



WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12P 7/62, C12N 9/88, 15/63, C07H 21/04		(11) International Publication Number: WO 97/22711 (43) International Publication Date: 26 June 1997 (26.06.97)	
(21) International Application Number: PCT/US96/2019		(81) Designated States: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(22) International Filing Date: 18 December 1996 (18.12.96)		Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.	
(30) Priority Data: 60/008,847 19 December 1995 (19.12.95) US			
(71) Applicant (for all designated States except US): REGENTS OF THE UNIVERSITY OF MINNESOTA (US/US), Morrill Hall, 100 Church Street S.E., Minneapolis, MN 55455 (US).			
(72) Inventors; and (75) Inventors/Applicants (for US only): SHERMAN, David, H. [US/US]; 2248 Drew Avenue South, St. Louis Park, MN 55416 (US); WILLIAMS, Mark, D. [US/US]; 1414 Breda Avenue, St. Paul, MN 55108 (US); XUE, Yongquan [CN/US]; 1403 Cleveland Avenue North, St. Paul, MN 55108 (US).			
(74) Agent: VIKSNINS, Ann, S.; Schwegman, Lundberg, Woessner & Kluth, P.O. Box 2938, Minneapolis, MN 55402 (US).			
(54) Title: METABOLIC ENGINEERING OF POLYHYDROXYALKANOATE MONOMER SYNTHASES			
(57) Abstract A novel pathway for the synthesis of polyhydroxyalkanoates is provided. A method of synthesizing a recombinant polyhydroxyalkanoate monomer synthase is also provided. These recombinant polyhydroxyalkanoate synthases are derived from multifunctional fatty acid synthases or polypeptide synthases and generate hydroxycyl acids, capable of polymerization by a polyhydroxyalkanoate synthase.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MI	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SG	Singapore
CI	Cote d'Ivoire	LI	Liechtenstein	SI	Slovenia
CN	China	LK	Sri Lanka	SK	Slovakia
CO	Colombia	LR	Liberia	SN	Senegal
CZ	Czech Republic	LT	Lithuania	SZ	Swaziland
DE	Germany	LV	Latvia	TD	Chad
DK	Denmark	LU	Luxembourg	TG	Togo
EE	Estonia	MC	Monaco	TJ	Tajikistan
ES	Spain	MD	Republic of Moldova	TT	Trinidad and Tobago
FI	Finland	MG	Madagascar	UA	Ukraine
FR	France	ML	Mali	UG	Uganda
GA	Gabon	MN	Mongolia	US	United States of America
		MR	Mauritania	UZ	Uzbekistan
				VN	Viet Nam

## METABOLIC ENGINEERING OF POLY- HYDROXYALKANOATE MONOMER SYNTHASES

2

### Background of the Invention

Polyhydroxyalkanoates (PHAs) are one class of biodegradable polymers. The first identified member of the PHAs thermoplastics was polyhydroxybutyrate (PHB), the polymeric ester of D(-)-3-hydroxybutyrate.

The biosynthetic pathway of PHB in the gram negative bacterium *Alcaligenes eutrophus* is depicted in Figure 1. PHAs related to PHB differ in the structure of the pendant arm, R (Figure 2). For example, R=CH<sub>3</sub> in PHB, while R=CH<sub>2</sub>CH<sub>3</sub> in polyhydroxyvalerate, and R=(CH<sub>2</sub>)<sub>6</sub>CH<sub>3</sub> in polyhydroxyoctanoate.

The genes responsible for PHB synthesis in *A. eutrophus* have been cloned and sequenced. (Peoples et al., *J. Biol. Chem.*, 264, 15293 (1989); Peoples et al., *J. Biol. Chem.*, 264, 15298 (1989)). Three enzymes:  $\beta$ -ketothiolase (*phbA*), acetoacetyl-CoA reductase (*phbB*), and PHB synthase (*phbC*) are involved in the conversion of acetyl-CoA to PHB. The PHB synthase gene encodes a protein of M<sub>r</sub>=63,900 which is active when introduced into *E. coli* (Peoples et al., *J. Biol. Chem.*, 264, 15298 (1989)).

Although PHB represents the archetypical form of a biodegradable thermoplastic, its physical properties preclude significant use of the homopolymer form. Pure PHB is highly crystalline and, thus, very brittle. However, unique physical properties resulting from the structural characteristics of the R groups in a PHA copolymer may result in a polymer with more desirable characteristics. These characteristics include altered crystallinity, UV weathering resistance, glass to rubber transition temperature (T<sub>g</sub>), melting temperature of the crystalline phase, rigidity and durability (Holmes et al., EPO 00052 459; Anderson et al., *Microbiol. Rev.*, 54, 450 (1990)). Thus, these polyesters behave as thermoplastics, with melting

temperatures of 50-180°C, which can be processed by conventional extension and molding equipment.

Traditional strategies for producing random PHA copolymers involve feeding short and long chain fatty acid monomers to bacterial cultures. However, this technology is limited by the monomer units which can be incorporated into a polymer by the endogenous PHA synthase and the expense of manufacturing PHAs by existing fermentation methods (Haywood et al., *EMS Microbiol. Lett.*, 52, 1 (1989); Poi et al., *Int. J. Biol. Macromol.*, 12, 106 (1990); Steinbuechel et al., In: *Novel Biomaterials from Biological Sources*, D. Byron (ed.), MacMillan, NY (1991); Valentin et al., *Appl. Microbiol. Biotechnol.*, 36, 507 (1992)).

The production of diverse hydroxyacyl-CoA monomers for homo- and co-polymeric PHAs also occurs in some bacteria through the reduction and condensation pathway of fatty acids. This pathway employs a fatty acid synthase (FAS) which condenses malonate and acetate. The resulting  $\beta$ -keto group undergoes three processing steps,  $\beta$ -keto reduction, dehydration, and enoyl reduction, to yield a fully saturated butyryl unit. However, this pathway provides only a limited array of PHA monomers which vary in alkyl chain length but not in the degree of alkyl group branching, saturation, or functionalization along the acyl chain.

The biosynthesis of polyketides, such as erythromycin, is mechanistically related to formation of long-chain fatty acids. However, polyketides, in contrast to FASs, retain ketone, hydroxyl, or olefinic functions and contain methyl or ethyl side groups interspersed along an acyl chain comparable in length to that of common fatty acids. This asymmetry in structure implies that the polyketide synthase (PKS), the enzyme system responsible for formation of these molecules, although mechanistically related to a

FAS, results in an end product that is structurally very different than that of a long chain fatty acid.

Because PHAs are biodegradable polymers that have the versatility to replace petrochemical-based thermoplastics, it is desirable that new, more economic methods be provided for the production of defined PHAs. Thus, what is needed are methods to produce recombinant PHA monomer synthases for the generation of PHA polymers.

#### 10 Summary of the Invention

The present invention provides a method of preparing a polyhydroxyalkanoate synthase. The method comprises introducing an expression cassette into a non-plant eukaryotic cell. The expression cassette comprises a DNA molecule encoding a polyhydroxyalkanoate synthase operably linked to a promoter functional in the non-plant eukaryotic cell. The DNA molecule encoding the polyhydroxyalkanoate synthase is then expressed in the cell. Thus, another embodiment of the invention provides a purified, isolated recombinant polyhydroxybutyrate synthase.

Another embodiment of the invention is a method of preparing a polyhydroxyalkanoate polymer. The method comprises introducing a first expression cassette and a second expression cassette into a eukaryotic cell. The first expression cassette comprises a DNA segment encoding a fatty acid synthase in which the dehydrase activity has been inactivated that is operably linked to a promoter functional in the eukaryotic cell. The second expression cassette comprises a DNA segment encoding a polyhydroxyalkanoate synthase operably linked to a promoter functional in the eukaryotic cell. The DNA segments in the expression cassettes are expressed in the cell so as to yield a polyhydroxyalkanoate polymer.

Another embodiment of the invention is a baculovirus expression cassette comprising a nucleic acid molecule encoding a polyhydroxyalkanoate synthase operably linked to a promoter functional in an insect cell.

5 The present invention also provides an expression cassette comprising a nucleic acid molecule encoding a polyhydroxyalkanoate monomer synthase operably linked to a promoter functional in a host cell. The nucleic acid molecule comprises a plurality of DNA segments. Thus, the nucleic acid molecule comprises at least a first and a second DNA segment. No more than one DNA segment is derived from the *eryA* gene cluster of *Saccharopolyspora erythraea*. The first DNA segment encodes a first module and the second DNA segment encodes a second module, wherein the DNA segments together encode a polyhydroxyalkanoate synthase.

Also provided is an isolated and purified DNA molecule. The DNA molecule comprises a plurality of DNA segments. Thus, the DNA molecule comprises at least a first and a second DNA segment. The first DNA segment encodes a first module and the second DNA segment encodes a second module. No more than one DNA segment is derived from the *eryA* gene cluster of *Saccharopolyspora erythraea*. Together the DNA segments encode a recombinant polyhydroxyalkanoate monomer synthase. A preferred embodiment of the invention employs a first DNA segment derived from the *veg* gene cluster of *Streptomyces*. Another preferred embodiment of the invention employs a second DNA segment derived from the *tyl* gene cluster of *Streptomyces*.

Yet another embodiment of the invention is a method of providing a polyhydroxyalkanoate monomer. The method comprises introducing a DNA molecule into a host cell. The DNA molecule comprises a DNA segment encoding a recombinant

5

polyhydroxyalkanoate monomer synthase operably linked to a promoter functional in the host cell. The DNA encoding the recombinant polyhydroxyalkanoate monomer synthase, which synthase comprises at least a first module and a second module, is expressed in the host cell so as to generate a polyhydroxyalkanoate monomer.

Also provided is a method of preparing a polyhydroxyalkanoate polymer. The method comprises introducing a first DNA molecule and a second DNA molecule into a host cell. The first DNA molecule comprises a DNA segment encoding a recombinant polyhydroxyalkanoate monomer synthase comprises a plurality of modules. Thus, the monomer synthase comprises at least a first module and a second module. The first DNA molecule is operably linked to a promoter functional in a host cell. The second DNA molecule comprises a DNA segment encoding a polyhydroxyalkanoate synthase operably linked to a promoter functional in the host cell. The DNAs encoding the recombinant polyhydroxyalkanoate monomer synthase and polyhydroxyalkanoate synthase are expressed in the host cell so as to generate a polyhydroxyalkanoate polymer.

Yet another embodiment of the invention is an isolated and purified DNA molecule. The DNA molecule comprises a plurality of DNA segments. That is, the DNA molecule comprises at least a first and a second DNA segment. The first DNA segment encodes a fatty acid synthase and the second DNA segment encodes a module of a polyketide synthase. A preferred embodiment of the invention employs a second DNA segment encoding a module which comprises a  $\beta$ -ketoacyl synthase amino-terminal to an acyltransferase

6

which is amino-terminal to a ketoreductase which is amino-terminal to an acyl carrier protein which is amino-terminal to a thioesterase.

The invention also provides a method of preparing a polyhydroxyalkanoate monomer. The method comprises introducing a DNA molecule comprising a plurality of DNA segments into a host cell. Thus, the DNA molecule comprises at least a first and a second DNA segment. The first DNA segment encodes a fatty acid synthase operably linked to a promoter functional in the host cell. The second DNA segment encodes a polyketide synthase. The second DNA segment is located 3' to the first DNA segment. The first DNA segment is linked to the second DNA segment so that the encoded protein is expressed as a fusion protein. The DNA molecule is then expressed in the host cell so as to generate a polyhydroxyalkanoate monomer.

Another embodiment of the invention is an expression cassette comprising a DNA molecule comprising a DNA segment encoding a fatty acid synthase and a polyhydroxyalkanoate synthase.

Also provided is a method of providing a polyhydroxyalkanoate monomer synthase. The method comprises introducing an expression cassette into a host cell. The expression cassette comprises a DNA molecule encoding a polyhydroxyalkanoate monomer synthase operably linked to a promoter functional in the host cell. The monomer synthase comprises a plurality of modules. Thus, the monomer synthase comprises at least a first and second module which together encode the monomer synthase.

A further embodiment of the invention is an isolated and purified DNA molecule comprising a DNA segment which encodes a *Streptomyces venezuelae* polyhydroxyalkanoate monomer synthase, a biologically active variant or subunit thereof. Preferably, the DNA segment encodes a polypeptide having an amino acid sequence comprising SEQ ID NO.2. Preferably, the DNA segment comprises

SEQ ID NO:1. The DNA molecules of the invention are double stranded or single stranded. A preferred embodiment of the invention is a DNA molecule that has at least about 70%, more preferably at least about 80%, and even more preferably at least about 90%, identity to the DNA segment comprising SEQ ID NO:1, e.g., a "variant" DNA molecule. A variant DNA molecule of the invention can be prepared by methods well known to the art, including oligonucleotide-mediated mutagenesis. See Adelman et al., DNA, 2, 183 (1983) and Sambrook et al., Molecular Cloning: A Laboratory Manual (1989).

The invention also provides an isolated, purified polyhydroxyalkanoate monomer synthase, e.g., a polypeptide having an amino acid sequence comprising SEQ ID NO:2, a biologically active subunit, or a biologically active variant thereof. Thus, the invention provides a variant polypeptide having at least about 80%, more preferably at least about 90%, and even more preferably at least about 95%, identity to the polypeptide having an amino acid sequence comprising SEQ ID NO:2. A preferred variant polypeptide, or subunit of a polypeptide, of the invention includes a variant or subunit polypeptide having at least about 10%, more preferably at least about 50% and even more preferably at least about 90%, the activity of the polypeptide having the amino acid sequence comprising SEQ ID NO:2. Preferably, a variant polypeptide of the invention has one or more conservative amino acid substitutions relative to the polypeptide having the amino acid sequence comprising SEQ ID NO:2. For example, conservative substitutions include aspartic-glutamic as acidic amino acids; lysine/arginine/histidine as basic amino acids; leucine/isoleucine, methionine/valine, alanine/valine as hydrophobic amino acids; serine/glycine/alanine/threonine as hydrophilic amino acids. The biological activity of a polypeptide of the invention can be measured by methods well known to the art.

As used herein, a "linker region" is an amino acid sequence present in a multifunctional protein which is less well conserved in amino acid sequence than an amino acid sequence with catalytic activity.

As used herein, an "extender unit" catalytic or enzymatic domain is an acyl transferase in a module that catalyzes chain elongation by adding 2-4 carbon units to an acyl chain and is located carboxy-terminal to another acyl transferase. For example, an extender unit with methylmalonylCoA specificity adds acyl groups to a methylmalonylCoA molecule.

As used herein, a "polyhydroxyalkanoate" or "PHA" polymer includes, but is not limited to, linked units of related, preferably heterologous, hydroxyalkanoates such as 3-hydroxybutyrate, 3-hydroxyvalerate, 3-hydroxycaproate, 3-hydroxyheptanoate, 3-hydroxyhexanoate, 3-hydroxyoctanoate, 3-hydroxyundecanoate, and 3-hydroxydodecanoate, and their 4-hydroxy and 5-hydroxy counterparts.

As used herein, a "Type I polyketide synthase" is a single polypeptide with a single set of iteratively used active sites. This is in contrast to a Type II polyketide synthase which employs active sites on a series of polypeptides.

As used herein, a "recombinant" nucleic acid or protein molecule is a molecule where the nucleic acid molecule which encodes the protein has been modified *in vitro*, so that its sequence is not naturally occurring, or corresponds to naturally occurring sequences that are not positioned as they would be positioned in a genome which has not been modified.

As used herein, a "multifunctional protein" is one where two or more enzymatic activities are present on a single polypeptide.

As used herein, a "module" is one of a series of repeated units in a multifunctional protein, such as a Type I polyketide synthase or a fatty acid synthase.

As used herein, a "premature termination product" is a product which is produced by a recombinant multifunctional protein which is different than the product produced by the non-recombinant multifunctional protein. In general, the product produced by the recombinant multifunctional protein has fewer acyl groups.

As used herein, a DNA that is "derived from" a gene cluster, is a DNA that has been isolated and purified *in vitro* from genomic DNA, or synthetically prepared on the basis of the sequence of genomic DNA.

As used herein, the terms "isolated and/or purified" refer to *in vitro* isolation of a DNA or polypeptide molecule from its natural cellular environment, and from association with other components of the cell, such as nucleic acid or polypeptide, so that it can be sequenced, replicated and/or expressed. Moreover, the DNA may encode more than one recombinant Type I polyketide synthase and/or fatty acid synthase. For example, "an isolated DNA molecule encoding a polyhydroxyalkanoate monomer synthase" is RNA or DNA containing greater than 7, preferably 15, and more preferably 20 or more sequential nucleotide bases that encode a biologically active polypeptide, fragment, or variant thereof, that is complementary to the non-coding, or complementary to the coding strand, of a polyhydroxyalkanoate monomer synthase RNA, or hybridizes to the RNA or DNA encoding the polyhydroxyalkanoate monomer synthase and remains stably bound under stringent conditions, as defined by methods well known to the art, e.g., in Sambrook et al., *supra*.

### Brief Description of the Figures

Figure 1: The PHB biosynthetic pathway in *A. eutrophus*.

*eutrophus*.

Figure 2: Molecular structure of common bacterial PHAs. Most of the known PHAs are polymers of 3-hydroxy acids possessing the general formula shown. For example,  $R=CH_3$  in PHB,  $R=CH_2CH_3$  in polyhydroxyvalerate (PHV), and  $R=(CH_2)_4CH_3$  in polyhydroxyoctanoate (PHO).

Figure 3: Comparison of the natural and recombinant pathways for PHB synthesis. The three enzymatic steps of PHB synthesis in bacteria involving 3-ketothiolase, acetoacetyl-CoA reductase, and PHB synthase are shown on the left. The two enzymatic steps involved in PHB synthesis in the pathway in *S21* cells containing a rat fatty acid synthase with an inactivated dehydrase domain (ratFAS206) are shown on the right.

Figure 4: Schematic diagram of the molecular organization of the *tyl* polyketide synthase (PKS) gene cluster. Open arrows correspond to individual open reading frames (ORFs) and numbers above an ORF denote a multifunctional module or synthase unit (SU). AT=acyltransferase; ACP=acyl carrier protein; KS= $\beta$ -ketoacyl synthase; KR=ketoreductase; DH=dehydrase; ER=enoyl reductase; TE=thioesterase; MM=methylmalonyl-CoA; M=malonyl CoA; EM=ethylmalonyl CoA. Module 7 in *tyl* is also known as Module F.

Figure 5: Schematic diagram of the molecular organization of the *met* PKS gene cluster.

Figure 6: Strategy for producing a recombinant PHA monomer synthase by domain replacement.

Figure 7: (A) 10% SDS-PAGE gel showing samples from various stages of the purification of PHA synthase; lane 1,

11

molecular weight markers; lane 2, total protein of uninfected insect cells; lane 3, total protein of insect cells expressing a rat FAS (200 kDa; Joshi et al., *Biochem J.* 296, 143 (1993)); lane 4, total protein of insect cells expressing PHA synthase; lane 5, soluble protein from sample in lane 4; lane 6, pooled hydroxylapatite (HA) fractions containing PHA synthase. (B) Western analysis of an identical gel using rabbit- $\alpha$ -PHA synthase antibody as probe. Bands designated with arrows are: a, intact PHB synthase with N-terminal alanine at residue 7 and serine at residue 10 (A7/S10); b, 44 kDa fragment of PHB synthase with N-terminal alanine at residue 181 and asparagine at residue 185 (A181/N185); c, PHB synthase fragment of approximately 30 kDa apparently blocked based on resistance to Edman degradation; d, 22 kDa fragment with N-terminal glycine at residue 187 (G187). Band d apparently does not react with rabbit- $\alpha$ -PHB synthase antibody (B, lane 6). The band of similar size in B, lane 4 was not further identified.

Figure 8: N-terminal analysis of PHA synthase purified from insect cells. (a) The expected N-terminal 25 amino acid sequence of *A. eutrophus* PHA synthase. (b&c) The two N-terminal sequences determined for the *A. eutrophus* PHA synthase produced in insect cells. The bolded sequences are the actual N-termini determined.

Figure 9: Spectrophotometric scans of substrate, 3-hydroxybutyrate CoA (HBCoA) and product, CoA. The wavelength at which the direct spectrophotometric assays were carried out (232 nm) is denoted by the arrow; substrate, HBCoA (•) and product, CoA (°).

Figure 10: Velocity of the hydrolysis of HBCoA as a function of substrate concentration. Assays were carried out in 40 or 200  $\mu$ l assay volumes with enzyme concentration remaining constant at 0.95 mg/ml (3.8  $\mu$ g/40 $\mu$ l assay). Velocities were calculated from the linear portions of the assay curves subsequent to the characteristic lag

12

period. The substrate concentration at half-optimal velocity, the apparent  $K_m$  value, was estimated to be 2.5 mM from this data.

Figure 11: Double reciprocal plot of velocity versus substrate concentration. The concave upward shape of this plot is similar to results obtained by Fukui et al. (*Arch. Microbiol.* 110, 149 (1976)) with granular PHA synthase from *Z. ramigera*.

Figure 12: Velocity of the hydrolysis of HBCoA as a function of enzyme concentration. Assays were carried out in 40  $\mu$ l assay volumes with the concentration HBCoA remaining constant at 8  $\mu$ M.

Figure 13: Specific activity of PHA synthase as a function of enzyme concentration.

Figure 14: pH activity curve for soluble PHA synthase produced using the baculovirus system. Reactions were carried out in the presence of 200 mM  $P_i$ . Buffers of pH<10 were prepared with potassium phosphate, while buffers of pH>10 were prepared with the appropriate proportion of  $Na_2PO_4$ .

Figure 15: Assays of the hydrolysis of HBCoA with varying amounts of PHA synthase. Assays were carried out in 40  $\mu$ l assay volumes with the concentration of HBCoA remaining constant at 8  $\mu$ M. Initial  $A_{232}$  values, originally between 0.62 and 0.77, were normalized to 0.70. Enzyme amounts used in these assays were, from the upper-most curve, 0.38, 0.76, 1.14, 1.52, 1.90, 2.28, 2.66, 3.02, 3.42, 7.6, and 15.2  $\mu$ g, respectively.

Figure 16: SDS/PAGE analysis of proteins synthesized at various time-points during infection of S/21 cells. Approximately 0.5 mg of total cellular protein from various samples was fractionated on a 10% polyacrylamide gel. Samples include: uninfected cells, lanes 1-4, days 0, 1, 2, 3 respectively; infection with BacPAK6:phbC alone,

1 3

lanes 5-8, days 0, 1, 2, 3 respectively; infection with baculoviral clone containing ratFAS206 alone, lanes 9-12, days 0, 1, 2, 3 respectively; and ratFAS206 and BacPAK6 infected cells, lanes 13-16 days 0, 1, 2, 3, respectively. A=mobility of FAS, B=mobility of PHA synthase.

5 Molecular weight standard lanes are marked M.

Figure 17: Gas chromatographic evidence for PHB accumulation in *S/21* cells. Gas chromatograms from various samples are superimposed. PHB standard (Sigma) is chromatogram #7 showing a propylhydroxybutyrate elution time of 10.043 minutes (s, arrow). The gas chromatograms of extracts of the uninfected (#1); singly infected with ratFAS206 (#2, day 3); and singly infected with PHA synthase (#3, day 3) are shown at the bottom of the figure. Gas chromatograms of extracts of dual-infected cells at day 1 (#4), 2 (#5), and 3 (#6) are also shown exhibiting a peak eluting at 10.096 minutes (x, arrow). The peak of dual-infected, day 3 extract (#6) was used for mass spectrometry (MS) analysis.

Figure 18: Gas-chromatography-mass spectrometry analysis of PHB. The characteristic fragmentation of propylhydroxybutyrate at  $m/z$  of 43, 60, 87, and 131 is shown. A) standard PHB from bacteria (Sigma), and B) peak X from ratFAS206 and BacPAK6: phbC baculovirus infected, day 3 (#6, Figure 17) *S/21* cells expressing rat FAS dehydrase inactivated protein and PHA synthase.

Figure 19: Map of the *vep* (*Streptomyces venezuelae* polyene encoding) gene cluster.

Figure 20: Plasmid map of pDHS502.

Figure 21: Plasmid map of pDHS505.

Figure 22: Cloning protocol for pDHS505.

Figure 23: Nucleotide sequence (SEQ ID NO:1) and

30 corresponding amino acid sequence (SEQ ID NO:22) of the *vep* ORF.

1 4

### Detailed Description of the Invention

The invention described herein can be used for the production of a diverse range of biodegradable PHA polymers through genetic redesign of DNA encoding a FAS or *Streptomyces* spp. Type I PKS polypeptide to provide a recombinant PHA monomer synthase. Different PHA synthases can then be tested for their ability to polymerize the monomers produced by the recombinant PHA synthase into a biodegradable polymer. The invention also provides a method by which various PHA synthases can be tested for their specificity with respect to different monomer substrates.

The potential uses and applications of PHAs produced by PHA monomer synthases and PHA synthases includes both medical and industrial applications. Medical applications of PHAs include surgical pins, sutures, staples, swabs, wound dressings, blood vessel replacements, bone replacements and plates, stimulation of bone growth by piezoelectric properties, and biodegradable carrier for long-term dosage of pharmaceuticals. Industrial applications of PHAs include disposable items such as baby diapers, packaging containers, bottles, wrappings, bags, and films, and biodegradable carriers for long-term dosage of herbicides, fungicides, insecticides, or fertilizers.

In animals, the biosynthesis of fatty acids *de novo* from malonyl-CoA is catalyzed by FAS. For example, the rat FAS is a homodimer with a subunit structure consisting of 2505 amino acid residues having a molecular weight of 272,340 Da. Each subunit consists of seven catalytic activities in separate physical domains (Amy et al., *Proc. Natl. Acad. Sci. USA*, 86, 3114 (1989)). The physical location of six of the catalytic activities, ketoacyl synthase (KS), malonyl/acetyltransferase (M/AT), enoyl reductase (ER), ketoreductase (KR), acyl carrier protein (ACP), and thioesterase (TE), has been established by (1) the identification of the various active site residues within the overall amino acid sequence by isolation of



catalytically active fragments from limited proteolytic digests of the whole FAS, (2) the identification of regions within the FAS that exhibit sequence similarity with various monofunctional proteins, (3) expression of DNA encoding an amino acid sequence with catalytic activity to produce recombinant proteins, and (4) the identification of DNA that does not encode catalytic activity, i.e., DNA encoding a linker region. (Smith et al., *Proc. Natl. Acad. Sci. USA*, 73, 1184 (1976); Tsukamoto et al., *J. Biol. Chem.*, 263, 16225 (1988); Rangan et al., *J. Biol. Chem.*, 266, 19180 (1991)).

The seventh catalytic activity, dehydrase (DH), was identified as physically residing between AT and ER by an amino acid comparison of FAS with the amino acid sequences encoded by the three open reading frames of the *eryA* polyketide synthase (PKS) gene cluster of *Saccharopolyspora erythraea*. The three polypeptides that comprise this PKS are constructed from "modules" which resemble animal FAS, both in terms of their amino acid sequence and in the ordering of the constituent domains (Donadio et al., *Gene*, 111, 51 (1992); Benh et al., *Eur. J. Biochem.*, 204, 39 (1992)).

One embodiment of the invention employs a FAS in which the DH is inactivated (FAS DH-). The FAS DH- employed in this embodiment of the invention is preferably a eukaryotic FAS DH- and, more preferably, a mammalian FAS DH-. The most preferred embodiment of the invention is a FAS where the active site in the DH has been inactivated by mutation. For example, Joshi et al. (*J. Biol. Chem.*, 268, 22508 (1993)) changed the His<sup>878</sup> residue in the rat FAS to an alanine residue by site directed mutagenesis. *In vitro* studies showed that a FAS with this change (ratFAS206) produced 3-hydroxybutyryl-CoA as a premature termination product from acetyl-CoA, malonyl-CoA and NADPH.

As shown below, a FAS DH- effectively replaces the  $\beta$ -ketothiolase and acetoacetyl-CoA reductase activities of the natural pathway by producing D(-)-3-hydroxybutyrate as a premature termination product, rather than the usual 16-carbon product, palmitic acid. This premature termination product can then be incorporated into PHB by a PHB synthase (See Example 2).

Another embodiment of the invention employs a recombinant *Streptomyces* spp. PKS to produce a variety of  $\beta$ -hydroxyCoA esters that can serve as monomers for a PHA synthase. One example of a DNA encoding a Type I PKS is the *eryA* gene cluster, which governs the synthesis of erythromycin aglycone deoxerythronolide B (DEB). The gene cluster encodes six repeated units, termed modules or synthase units (SUs). Each module or SU, which comprises a series of putative FAS-like activities, is responsible for one of the six elongation cycles required for DEB formation. Thus, the processive synthesis of asymmetric acyl chains found in complex polyketides is accomplished through the use of a programmed protein template, where the nature of the chemical reactions occurring at each point is determined by the specificities in each SU.

Two other Type I PKS are encoded by the *tyl* (tylosin) (Figure 4) and *met* (methymycin) (Figure 5) gene clusters. The macrolide multifunctional synthases encoded by *tyl* and *met* provide a greater degree of metabolic diversity than that found in the *eryA* gene cluster. The PKSs encoded by the *eryA* gene cluster only catalyze chain elongation with methylmalonyl-CoA, as opposed to *tyl* and *met* PKSs, which catalyze chain elongation with malonyl-CoA, methylmalonyl-CoA and ethylmalonyl-CoA. Specifically, the *tyl* PKS includes two malonyl-CoA extender units and one ethylmalonyl-CoA extender unit, and the *met* PKS includes one malonyl-CoA extender unit. Thus, a preferred embodiment of the invention includes, but is

17

not limited to, replacing catalytic activities encoded in *met* PKS open reading frame 1 (ORF1) to provide a DNA encoding a protein that possesses the required keto group processing capacity and short chain acyl/CoA ester starter and extender unit specificity necessary to provide a saturated  $\beta$ -hydroxyhexanoyl/CoA or unsaturated  $\beta$ -hydroxyhexenoyl/CoA monomer.

In order to manipulate the catalytic specificities within each module, DNA encoding a catalytic activity must remain undisturbed. To identify the amino acid sequences between the amino acid sequences with catalytic activity, the "linker regions," amino acid sequences of related modules, preferably those encoded by more than one gene cluster, are compared. Linker regions are amino acid sequences which are less well conserved than amino acid sequences with catalytic activity. Witkowski et al., Eur. J. Biochem. 198, 571 (1991).

In an alternative embodiment of the invention, to provide a DNA encoding a Type I PKS module with a TE and lacking a functional DH, a DNA encoding a module F, containing KS, MT, KR, ACP, and TE catalytic activities, is introduced at the 3' end of a DNA encoding a first module (Figure 6). Module F introduces the final (R)-3-hydroxyl acyl group at the final step of PHA monomer synthesis, as a result of the presence of a TE domain. DNA encoding a module F is not present in the *eryA* PKS gene cluster (Donadio et al., *supra*, 1991).

A DNA encoding a recombinant monomer synthase is inserted into an expression vector. The expression vector employed varies depending on the host cell to be transformed with the expression vector. That is, vectors are employed with transcription, translation and/or post-translational signals, such as targeting signals, necessary for efficient expression of the genes in various host cells

18

into which the vectors are introduced. Such vectors are constructed and transformed into host cells by methods well known in the art. See Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor (1989). Preferred host cells for the vectors of the invention include insect, bacterial, and plant cells. Preferred insect cells include *Spodoptera frugiperda* cells such as Sf21, and *Trichoplusia ni* cells. Preferred bacterial cells include *Escherichia coli*, *Streptomyces* and *Pseudomonas*. Preferred plant cells include monocot and dicot cells, such as maize, rice, wheat, tobacco, legumes, carrot, squash, canola, soybean, potato, and the like.

Moreover, the appropriate subcellular compartment in which to locate the enzyme in eukaryotic cells must be considered when constructing eukaryotic expression vectors. Two factors are important: the site of production of the acetyl-CoA substrate, and the available space for storage of the PHA polymer. To direct the enzyme to a particular subcellular location, targeting sequences may be added to the sequences encoding the recombinant molecules.

The baculovirus system is particularly amenable to the introduction of DNA encoding a recombinant FAS or a PKS monomer synthase because an increasing variety of transfer plasmids are becoming available which can accommodate a large insert, and the virus can be propagated to high titers. Moreover, insect cells are adapted readily to suspension culture, facilitating relatively large scale recombinant protein production. Further, recombinant proteins tend to be produced exclusively as soluble proteins in insect cells, thus, obviating the need for refolding, a task that might be particularly daunting in the case of a large multifunctional protein. The Sf21/baculovirus system has routinely expressed milligram quantities of catalytically active recombinant fatty acid synthase. Finally, the baculovirus/insect cell system provides the ability to construct and

analyze different synthase proteins for the ability to polymerize monomers into unique biodegradable polymers.

A further embodiment of the invention is the introduction of at least one DNA encoding a PHA synthase and a DNA encoding a PHA monomer synthase into a host cell. Such synthases include, but are not limited to, *A. eutrophus* 3-hydroxy, 4-hydroxy, and 5-hydroxy alkanolate synthases, *Rhodococcus ruber* C<sub>3</sub>-C<sub>5</sub> hydroxyalkanoate synthases, *Pseudomonas oleovorans* C<sub>6</sub>-C<sub>14</sub> hydroxyalkanoate synthases, *P. putida* C<sub>6</sub>-C<sub>14</sub> hydroxyalkanoate synthases, *P. aeruginosa* C<sub>5</sub>-C<sub>10</sub> hydroxyalkanoate synthases, *P. resinovorans* C<sub>4</sub>-C<sub>10</sub> hydroxyalkanoate synthases, *Rhodospirillum rubrum* C<sub>4</sub>-C<sub>7</sub> hydroxyalkanoate synthases, *R. gelatinosus* C<sub>4</sub>-C<sub>7</sub>, *Thiobacillus pfennigii* C<sub>4</sub>-C<sub>8</sub> hydroxyalkanoate synthases, and *Bacillus megaterium* C<sub>4</sub>-C<sub>5</sub> hydroxyalkanoate synthases.

The introduction of DNA(s) encoding more than one PHA synthase may be necessary to produce a particular PHA polymer due to the specificities exhibited by different PHA synthases. As multifunctional proteins are altered to produce unusual monomeric structures, synthase specificity may be problematic for particular substrates. Although the *A. eutrophus* PHB synthase utilizes only C<sub>4</sub> and C<sub>5</sub> compounds as substrates, it appears to be a good prototype synthase for initial studies since it is known to be capable of producing copolymers of 3-hydroxybutyrate and 4-hydroxybutyrate (Kunioka et al., Macromolecules, 22, 694 (1989)) as well as copolymers of 3-hydroxyvalerate, 3-hydroxybutyrate, and 5-hydroxyvalerate (Doi et al., Macromolecules, 19, 2860 (1986)). Other synthases, especially those of *Pseudomonas aeruginosa* (Timm et al., Eur. J. Biochem., 209, 15 (1992)) and *Rhodococcus ruber* (Pieper et al., FEMS Microbiol. Lett., 96, 73 (1992)), can also be employed in the practice of the invention. Synthase specificity may be alterable through molecular biological methods.

In yet another embodiment of the invention, a DNA encoding a FAS and a PHA synthase can be introduced into a single expression vector, obviating the need to introduce the genes into a host cell individually.

A further embodiment of the invention is the generation of a DNA encoding a recombinant multifunctional protein, which comprises a FAS, of either eukaryotic or prokaryotic origin, and a PKS module F. Module F will carry out the final chain extension to include two additional carbons and the reduction of the  $\beta$ -keto group, which results in a (R)-3-hydroxy acyl CoA moiety.

To produce this recombinant protein, DNA encoding the FAS TE is replaced with a DNA encoding a linker region which is normally found in the ACP-KS interdomain region of bimodular ORFs. DNA encoding a module F is then inserted 3' to the DNA encoding the linker region. Different linker regions, such as those described below, which vary in length and amino acid composition, can be tested to determine which linker most efficiently mediates or allows the required transfer of the nascent saturated fatty acid intermediate to module F for the final chain elongation and keto reduction steps. The resulting DNA encoding the protein can then be tested for expression of long chain  $\beta$ -hydroxy fatty acids in insect cells, such as *Sf21* cells, or *Streptomyces*, or *Pseudomonas*. The expected 3-hydroxy C-18 fatty acid can serve as a potential substrate for PHA synthases which are able to accept long chain alkyl groups. A preferred embodiment of the invention is a FAS that has a chain length specificity between 4-22 carbons.

Examples of linker regions that can be employed in this embodiment of the invention include, but are not limited to, the ACP-KS linker regions encoded by the *tyl* ORF1 (ACP1-KS<sub>2</sub>; ACP2-

21

KS<sub>3</sub>), and ORF3 (ACP<sub>5</sub>-KS<sub>6</sub>), and *eryA* ORF1 (ACP<sub>1</sub>-KS<sub>1</sub>; ACP<sub>2</sub>-KS<sub>2</sub>), ORF2 (ACP<sub>3</sub>-KS<sub>4</sub>) and ORF3 (ACP<sub>5</sub>-KS<sub>6</sub>).

This approach can also be used to produce shorter chain fatty acid groups by limiting the ability of the FAS unit to generate long chain fatty acids. Mutagenesis of DNA encoding various FAS catalytic activities, starting with the KS, may result in the synthesis of short chain (R)-3-hydroxy fatty acids.

The PHA polymers are then recovered from the biomass. Large scale solvent extraction can be used, but is expensive. An alternative method involving heat shock with subsequent enzymatic and detergent digestive processes is also available (Byron, Trends Biotechnical, 5, 246 (1987); Holmes, In: Developments in Crystalline Polymers, D.C. Bassett (ed), pp. 1-65 (1988)). PHB and other PHAs are readily extracted from microorganisms by chlorinated hydrocarbons. Refluxing with chloroform has been extensively used; the resulting solution is filtered to remove debris and concentrated, and the polymer is precipitated with methanol or ethanol, leaving low-molecular-weight lipids in solution. Longer-side-chain PHAs show a less restricted solubility than PHB and are, for example, soluble in acetone. Other strategies adopted include the use of ethylene carbonate and propylene carbonate as disclosed by Lafferty et al. (Chem. Rundschau, 30, 14 (1977)) to extract PHB from biomass. Scandola et al., (Int. J. Biol. Microbiol., 10, 373 (1988)) reported that 1 M HCl-chloroform extraction of *Rhizobium meliloti* yielded PHB of M<sub>w</sub> = 6 x 10<sup>4</sup> compared with 1.4 x 10<sup>6</sup> when acetone was used.

Methods are well known in the art for the determination of the PHB or PHA content of microorganisms, the composition of PHAs, and the distribution of the monomer units in the polymer. Gas chromatography and high-pressure liquid chromatography are widely used for quantitative PHB analysis. See Anderson et al., Microbiol. Rev., 54, 450 (1990) for a review of such methods. NMR

22

techniques can also be used to determine polymer composition, and the distribution of monomer units.

The invention has been described with reference to various specific and preferred embodiments and will be further described by reference to the following detailed examples. It is understood however, that there are many extensive variations, and modifications on the basic theme of the present invention beyond that shown in the examples and description, which are within the spirit and scope of the present invention.

## I. Experimental Procedures

### Materials and Methods

**15 Materials.** Sodium R-(-)-3-hydroxybutyrate, coenzyme-A, ethylchloroformate, pyridine and diethyl ether were purchased from Sigma Chemical Co. Amberlite IR-120 was purchased from Mallinckrodt Inc. 6-O-(N-Heptylcarbamoyl)methyl  $\alpha$ -D-glucopyranoside (Hecameg) was obtained from Vegatec (Villeguif, France). Two-piece spectrophotometer cells with pathlengths of 0.1 (#20/0-Q-1) and 0.01 cm (#20/0-Q-0.1) were obtained from Starna Cells Inc., (Atascadero, CA). Rabbit anti-*A. eutrophus* PHA synthase antibody was a gracious gift from Dr. F. Sienne and S. Stoup (Biological Process Technology Institute, University of Minnesota), SZ21 cells and

**20** *T. ni* cells were kindly provided by Greg Franzen (R&D Systems, Minneapolis, MN) and Stephen Harsch (Department of Veterinary Pathobiology, University of Minnesota), respectively.

Plasmid pFAS206 and a recombinant baculoviral clone encoding FAS206 (Joshi et al., J. Biol. Chem., 268, 22508 (1993)) were generous gifts of A. Joshi and S. Smith. Plasmid pAc41 (Peoples et al.,

J. Biol. Chem., **264**, 15298, (1989)), the source of the *A. cutroplus* PHB synthase, was obtained from A. Sinskey. Baculovirus transfer vector, pBacPAK9, and linearized baculoviral DNA, were obtained from Clontech Inc. (Palo Alto, CA). Restriction enzymes, T4 DNA ligase, *E. coli* DH5 $\alpha$  competent cells, molecular weight standards, lipofectin reagent, Grace's insect cell medium, fetal bovine serum (FBS), and antibiotic/antimycotic reagent were obtained from GIBCO-BRL (Grand Island, NY). Tissue culture dishes were obtained from Corning Inc. Spinner flasks were obtained from Belco Glass Inc. Seaplaque agarose GTG was obtained from FMC Bioproducts Inc.

#### Methods.

**Preparation of R-3-HBCoA.** R-(-)-3 HBCoA was prepared by the mixed anhydride method described by Haywood et al., JEMS Microbiol. Lett., **52**, 1 (1989). 60 mg (0.58 mmol) of R-(-)-3 hydroxybutyric acid was freeze dried and added to a solution of 72 mg of pyridine in 10 ml diethyl ether at 0°C. Ethylchloroformate (100 mg) was added, and the mixture was allowed to stand at 4°C for 60 minutes. Insoluble pyridine hydrochloride was removed by centrifugation. The resulting anhydride was added, dropwise with mixing, to a solution of 100 mg coenzyme-A (0.13 mmol) in 4 ml 0.2 M potassium bicarbonate, pH 8.0 at 0°C. The reaction was monitored by the nitroprusside test of Stadtman, Meth. Enzymol., **3**, 931 (1957), to ensure sufficient anhydride was added to esterify all the coenzyme-A. The concentration of R-3-HBCoA was determined by measuring the absorbance at 260 nm ( $\epsilon = 16.8 \text{ mM}^{-1} \text{ cm}^{-1}$ ; 18).

**Construction of pBP-phbC.** The *phbC* gene (approximately 1.8 kb) was excised from pAet41 (Peoples et al., J. Biol. Chem., **264**, 15293 (1989)) by digestion with *Bst*BI and *Sst*II, purified as described by Williams et al. (Gene, **109**, 445 (1991)), and ligated to pBacPAK9

digested with *Bst*BI and *Sst*II. This resulted in pBP-phbC, the baculovirus transfer vector used in formation of recombinant baculovirus particles carrying *phbC*.

#### Large scale expression of PHA synthase. A 1 L culture of T.

*ni* cells ( $1.2 \times 10^6$  cells/ml) in logarithmic growth was infected by the addition of 50 ml recombinant viral stock solution ( $2.5 \times 10^8$  pfu/ml) resulting in a multiplicity of infection (MOI) of 10. This infected culture was split between two Belco spinners (350 ml/500 ml spinner, 700 ml/1 L spinner) to facilitate oxygenation of the culture. These cultures were incubated at 28°C and stirred at 60 rpm for 60 hours. Infected cells were harvested by centrifugation at  $1000 \times g$  for 10 minutes at 4°C. Cells were flash-frozen in liquid N<sub>2</sub> and stored in 4 equal aliquots, at -80°C until purification.

#### Insect cell maintenance and recombinant baculovirus

**formation.** *Sf*21 cells were maintained at 26-28°C in Grace's insect cell medium supplemented with 10% FBS, 1.0% pluronic F68, and 1.0% antibiotic/antimycotic (GIBCO-BRL). Cells were typically maintained in suspension at  $0.2 - 2.0 \times 10^6$ /ml in 60 ml total culture volume in 100 ml spinner flasks at 55-65 rpm. Cell viability during the culture period was typically 95-100%. The procedures for use of the transfer vector and baculovirus were essentially those described by the manufacturer (Clontech, Inc.). Purified pBP-phbC and linearized baculovirus DNA were used for cotransfection of *Sf*21 cells using the liposome mediated method (Felgner et al., Proc. Natl. Acad. Sci. USA, **84**, 7413 (1987)) utilizing Lipofectin (GIBCO-BRL). Four days later cotransfection supernatants were utilized for plaque purification. Recombinant viral clones were purified from plaque assay plates containing 1.5% Seaplaque GTG after 5-7 days at 28°C. Recombinant viral clone stocks were then amplified in T25-flask cultures (4 ml,  $3 \times 10^6$ /ml on day 0) for 4 days; infected cells were determined by their morphology and size and then screened by SDS/PAGE using 10%

2.5

polyacrylamide gels (Laemmli, *Nature*, 222, 680 (1970)) for production of PHA synthase.

#### Purification of PHA synthase from BTI-TN-5B1-4 *T. ni* cells.

Purification of PHA synthase was performed according to the method of Gemgross et al., *Biochemistry*, 33, 9311 (1994) with the following alterations. One aliquot (110 mg protein) of frozen cells was thawed on ice and resuspended in 10 mM KPi (pH 7.2), 5% glycerol, and 0.05% Hecaneg (Buffer A) containing the following protease inhibitors at the indicated final concentrations: benzamide (2 mM), phenylmethylsulfonyl fluoride (PMSF, 0.4 mM), pepstatin (2 mg/ml), leupeptin (2.5 mg/ml), and Na-p-tosyl-L-lysine chloromethyl ketone (TLCK, 2 mM). EDTA was omitted at this stage due to its incompatibility with hydroxylapatite (HA). This mixture was homogenized with three series of 10 strokes each in two Thomas homogenizers while partially submerged in an ice bath and then sonicated for 2 minutes in a Branson Sonifier 250 at 30% cycle, 30% power while on ice. All subsequent procedures were carried out at 4°C.

The lysate was immediately centrifuged at 100000 x g in a Beckman 50.2Ti rotor for 80 minutes, and the resulting supernatant (10.5 ml, 47 mg) was immediately filtered through a 0.45 mm Uniflow filter (Schleicher and Schuell Inc., Keene, N.H.) to remove any remaining insoluble matter. Aliquots of the soluble fraction (1.5 ml, 7 mg) were loaded onto a 5 ml BioRad Econo-Pac HTP column that had been equilibrated with Buffer A (+ protease inhibitor mix) attached to a BioRad Econo-system, and the column was washed with 30 ml Buffer A. All chromatographic steps were carried out at a flow rate of 0.8 ml/minute. PHA synthase was eluted from the HA column with a 32 x 32 ml linear gradient from 10 to 300 mM KPi.

Fraction collection tubes were prepared by addition of 30 ml of 100 mM EDTA to provide a metalloprotease inhibitor at 1 mM

2.6

immediately after HA chromatography. PHA synthase was eluted in a broad peak between 110-180 mM KPi. Fractions (3 ml) containing significant PHA synthase activity were pooled and stored at 0°C until the entire soluble fraction had been run through the chromatographic process. Pooled fractions then were concentrated at 4°C by use of a Centriprep-30 concentrator (Amicon) to 3.8 mg/ml. Aliquots (0.5 ml) were either flash-frozen and stored in liquid N<sub>2</sub> or glycerol was added to a final concentration of 50% and samples (1.9 mg/ml) were stored at -20°C.

#### Western analysis.

Samples of *T. ni* cells were fractionated by SDS-PAGE on 10% polyacrylamide gels, and the proteins then were transferred to 0.2 mm nitrocellulose membranes using a BioRad Transblot SD Semi-Dry electrophoretic transfer cell according to the manufacturer. Proteins were transferred for 1 hour at 15 V. The membrane was rinsed with doubly distilled H<sub>2</sub>O, dried, and treated with phosphate-buffered saline (PBS) containing 0.05% Tween-20 (PBS-Tween) and 3% nonfat dry milk to block non-specific binding sites. Primary antibody (rabbit anti-PHA synthase) was applied in fresh blocking solution and incubated at 25°C for 2 hours. Membranes were then washed four times for 10 minutes with PBS-Tween followed by the addition of horseradish peroxidase-conjugated goat-anti-rabbit antibody (Boehringer-Mannheim) diluted 10,000X in fresh blocking solution and incubated at 25°C for 1 hour. Membranes were washed finally in three changes (10 minutes) of PBS, and the immobilized peroxidase label was detected using the chemiluminescent LumiGLO substrate kit (Kirkegaard and Perry, Gaithersburg, MD) and X-ray film.

#### N-terminal analysis.

Approximately 10 mg of purified PHA synthase was run on a 10% SDS-polyacrylamide gel, transferred to PVDF (Immobilion-PSQ, Millipore Corporation, Bedford, MA), stained with Amido Black, and sequenced on a 494 Protein

Sequencer (Perkin-Elmer, Applied Biosystems Division, Foster City, California).

**Double-infection protocol.** Four 100 ml spinner flasks were each inoculated with  $8 \times 10^7$  cells in 50 ml of fresh insect medium. To flask 1, an additional 20 ml of fresh insect medium was added (uninfected control); to flask 2, 10 ml BacPAK6::phbC viral stock ( $1 \times 10^8$  pfu/ml) and 10 ml fresh insect medium were added; to flask 3, 10 ml BacPAK6::FAS206 viral stock ( $1 \times 10^8$  pfu/ml) and 10 ml fresh insect medium were added; and to flask 4, 10 ml BacPAK6::phbC viral stock ( $1 \times 10^8$  pfu/ml) and 10 ml BacPAK6::FAS206 viral stock ( $1 \times 10^8$  pfu/ml) were added. These viral infections were carried out at a multiplicity of infection of approximately 10. Cultures were maintained under normal growth conditions and 15 ml samples were removed at 24, 48, and 72 hour time points. Cells were collected by gentle centrifugation at  $1000 \times g$  for 5 minutes, the medium was discarded, and the cells were immediately stored at  $-70^\circ\text{C}$ .

**PHA synthase assays.** Coenzyme A released by PHA synthase in the process of polymerization was monitored precisely as described by Gerngross et al. (*supra*) using 5,5'-dithiobis (2-nitrobenzoic acid, DTNB) (Ellman, Arch. Biochem. Biophys., **82**, 70 (1959)).

The presence of HBCoA was monitored spectrophotometrically. Assays were performed at  $25^\circ\text{C}$  in a Hewlett Packard 8452A diode array spectrophotometer equipped with a water jacketed cell holder. Two-piece Starna Spectrosil spectrophotometer cells with pathlengths of 0.1 and 0.01 cm were employed to avoid errors arising from the compression of the absorbance scale at higher values. Absorbance was monitored at 232 nm, and  $E_{232\text{nm}}$  of  $4.5 \times 10^3$  M $^{-1}$  cm $^{-1}$  was used in calculations. One unit (U) of enzyme is the amount required to hydrolyze 1 mmol of substrate minute $^{-1}$ . Buffer

(0.15 M KPi, pH 7.2) and substrate were equilibrated to  $25^\circ\text{C}$  and then combined in an Eppendorf tube also at  $25^\circ\text{C}$ . Enzyme was added and mixed once in the pipet tip used to transfer the entire mixture to the spectrophotometer cell. The two piece cell was immediately assembled, placed in the spectrophotometer with the cell holder (type CH) adapted for the standard 10 mm path length cell holder of the spectrophotometer. Manipulations of sample, from mixing to initiation of monitoring, took only 10-15 seconds. Absorbance was continually monitored for up to 10 minutes. Calibration of reactions was against a solution of buffer and enzyme (no substrate) which lead to absorbance values that represented substrate only.

**PHB assay.** PHB was assayed from S21 cell samples according to the propanolysis method of Riis et al., J. Chrom., **445**, 285 (1988). Cell pellets were thawed on ice, resuspended in 1 ml cold ddH<sub>2</sub>O and transferred to 5 ml screwtop test tubes with teflon seals. 2 ml ddH<sub>2</sub>O was added, the cells were washed and centrifuged and then 3 ml of acetone were added and the cells washed and centrifuged. The samples were then desiccated by placing them in a  $94^\circ\text{C}$  oven for 12 hours. The following day 0.5 ml of 1,2-dichloroethane, 0.5 ml acidified propanol (20 ml HCl, 80 ml 1-propanol) and 50 ml benzoic acid standard were added and the sealed tubes were heated to  $100^\circ\text{C}$  in a boiling water bath for 2 hours with periodic vortexing. The tubes were cooled to room temperature and the organic phase was used for gas-chromatographic (GC) analysis using a Hewlett Packard 5890A gas-chromatograph equipped with a Hewlett Packard 7673A automatic injector and a fused silica capillary column, DB-WAX 30W of 30 meter length. Positive samples were further subjected to GC-mass spectrometric (MS) analysis for the presence of propylhydroxybutyrate using a Kratos MS25 GC/MS. The following parameters were used: source temperature,  $210^\circ\text{C}$ ; voltage, 70eV; and accelerating voltage, 4 KeV.

29

#### Catalytic activities.

Ketoacyl synthase (KS) activity was assessed radiochemically by the condensation- $^{14}\text{C}$ CO<sub>2</sub> exchange reaction (Smith et al., PNAS, USA, 73, 1184 (1976)).

Transferase (AT) activity was assayed, using malonyl-CoA as donor and pantetheine as acceptor, by determining spectrophotometrically the free CoA released in a coupled ATP citrate-lyase-malate dehydrogenase reaction (see, Rangen et al., J. Biol. Chem., 266, 19180 (1991)).

Ketoreductase (KR) was assayed spectrophotometrically at 340 nm: assay systems contained 0.1 M potassium phosphate buffer (pH 7), 0.15 mM NADPH, enzyme and either 10 mM *trans*-1-decalone or 0.1 mM acetoacetyl-CoA substrate.

Dehydrase (DH) activity was assayed spectrophotometrically at 270 nm using S-DL- $\beta$ -hydroxybutyryl N-acetylcysteamine as substrate (Kumar et al., J. Biol. Chem., 245, 4732 (1970)).

Fatty acid reductase (ER) activity was assayed spectrophotometrically at 340 nm essentially as described by Strometal (J. Biol. Chem., 254, 8159 (1979)): the assay system contained 0.1 M potassium phosphate buffer (pH 7), 0.15 mM NADPH, 0.375 mM crotonoyl-CoA, 20  $\mu\text{M}$  CoA and enzyme.

Thioesterase (TE) activity was assessed radiochemically by extracting and assaying the [ $^{14}\text{C}$ ]palmitic acid formed from [ $^{14}\text{C}$ ]palmitoyl-CoA during a 3 minute incubation (Smith, Meth. Enzymol., 71, 181 (1981)), the assay was in a final volume of 0.1 ml, 25 mM potassium phosphate buffer (pH 8), 20  $\mu\text{M}$  [ $^{14}\text{C}$ ]palmitoyl-CoA (20 nCi) and enzyme.

Assay of overall fatty acid synthase activity was performed spectrophotometrically as described previously by Smith et al. (Meth.

30

Enzymol., 35, 65 (1975)). All enzyme activities were assayed at 37°C except the transferase, which was assayed at 20°C. Activity units indicate nmol of substrate consumed/minute. All assays were conducted, at a minimum, at two different protein concentrations with the appropriate enzyme and substrate blanks included.

#### Example 1

Expression of *A. eutrophus* PHA synthase using a baculovirus system.

Recent work has shown that PHA synthase from *A. eutrophus* can be overexpressed in *E. coli*, in the absence of 3-ketothiolase and acetoacetyl-CoA reductase (Germgross et al. *supra*) and can be expressed in plants (See Poirier et al., Biotect., 13, 142 (1995) for a review). Isolation of the soluble form of PHA synthase provides opportunities to examine the mechanistic details of the priming and initiation reactions. Because the baculovirus system has been successful for the expression of a number of prokaryotic genes as soluble proteins, and insect cells, unlike bacterial expression systems, carry out a wide array of posttranslational modifications, the baculovirus expression system appeared ideal for the expression of large quantities of soluble PHA synthase, a protein that must be modified by phosphopantetheine in order to be catalytically active (Germgross et al., *supra*).

Purification of PHA synthase. The purification procedure employed for PHA synthase is a modification of Germgross et al. (*supra*) involving the elimination of the second liquid chromatographic step and inclusion of a protease-inhibitor cocktail in all buffers. All steps were carried out on ice or at 4°C except where noted. Frozen cells were thawed on ice in 10 ml of Buffer A (10 mM KPi, pH 7.2, 0.5% glycerol, and 0.05% Hecameg) and then immediately homogenized prior to centrifugation and HA chromatography.



anti-PHA synthase antibody corroborated the results of the sequencing and indicated the presence of at least three bands that resulted from proteolysis of PHA synthase (Figure 7B, Lanes 4-6). The antibody was specific for PHA synthase since neither *T. ni* nor baculoviral proteins showed reactivity (Figure 7B, Lanes 2 and 3). N-terminal protein sequencing (Figure 8) showed directly that the 44 kDa (band b) and 32 kDa (band d) proteins were derived from PHA synthase (fragments beginning at A181/N185 and at C387, respectively). The 35-40 kDa (band c) protein gave low sequencing yields and may contain a blocked N-terminus. Inspection of Figure 7B suggests that most degradation occurs following cell disruption since the total protein sample for this gel (lane 4) was prepared by boiling intact cells directly in SDS sample buffer while the HA sample (lane 6) went through the purification procedure described above.

**Assay of Synthase Activity.** Due to the significant level of expression obtained using the baculovirus system, the synthase activity could be assayed spectrophotometrically by monitoring hydrolysis of the thioester bond at 232 nm, the wavelength at which there is a maximum decrease in absorbance upon hydrolysis. The difference between substrate (HBCoA) and product (CoA) at this wavelength is shown in Figure 9. Absorbance of HBCoA and CoA at 232 nm occurs at a trough between two well separated peaks. Assays were carried out at pH 7.2 for comparative analysis with previous studies (Gerngross et al., *supra*). Substrate (R-(-)-3-HBCoA) substrate for these studies was prepared using the mixed anhydride method (Haywood et al., *supra*), and its concentration was determined by measuring  $A_{260}$ . The short pathlength cells (0.1 cm and 0.01 cm) allowed use of relatively high reaction concentrations while conserving substrate and enzyme. Assay results showed an initial lag period of 60 seconds prior to the linear decrease in  $A_{232}$ , and velocities were determined from the slope of these linear regions of

The results of these efforts are summarized in Table 1. A prominent band at 64 kDa is visible in total, soluble, and HA eluate protein samples fractionated by SDS/PAGE (lanes 4, 5, and 6 of Figure 7, respectively). The initial specific activity of the isolated PHA synthase was 20-fold higher than previous attempts at expression and purification of this polypeptide. Approximately 1000 units of PHB synthase have been purified, based on calculations from the direct spectrophotometric assay detailed below, with an overall recovery of activity of 70%. The large proportion of synthase present in the membrane fraction, and the fact that over 90% of the initial activity was found in the soluble fraction, suggests either that the synthase in the membrane fraction is in an inactive form or that the direct assay is not applicable to the initial, 12 U/mg, crude extract.

Table 1: Purification of PHA Synthase

sample	total units	vol (mL)	protein (mg)	specific activity (mg/ml)	recovery
total	1430	11.5	113	9.8	100
protein					
soluble	1340	10.5	47	4.5	93
protein					
pooled	1020	7.9	30	3.8	71
HA					
fractions					

N-terminal sequencing of the 64 kDa protein confirmed its identity as PHA synthase (Figure 8). Two prominent N-termini, at amino acid residue 7 (alanine) and residue 10 (serine) were obtained in a 3:2 ratio. This heterogeneous N-terminus presumably is the result of aminopeptidase activity. Western analysis using a rabbit-

3 3

the assay curves. The length of the lag period was variable and was inversely related to enzyme concentration. These data are consistent with those using PHA synthase purified from *E. coli* (Gemgross et al., *supra*).

Figures 10 and 11 show the V versus S and 1/V versus 1/S plots, respectively. The double reciprocal plot was concave upward which is similar to results obtained from studies of the granular PHA synthase from *Zooglea ramigera* (Fukui et al., *Arch. Microbiol.*, **110**, 149 (1976)) and suggests a complex reaction mechanism. Examinations of velocity and specific activity as a function of enzyme concentration are shown in Figures 12 and 13. These results confirm that specific activity of the synthase depends upon enzyme concentration. The pH activity curve for *A. eutrophus* PHA synthase purified from *T. ni* cells is shown in Figure 14. The curve shows a broad activity maximum centered around pH 8.5. This result agrees well with prior work on the *A. eutrophus* PHB synthase although it is significantly different than results obtained for the PHB synthase from *Z. ramigera* for which the optimum was determined to be pH 7.0.

The effect of varying enzyme concentration in the presence of a fixed amount of substrate revealed an intriguing trend (Figure 15). From these data it appears that the extent of polymerization is dependent on the amount of enzyme included in the reaction mixture. This could be explained if there is a "terminal length" limitation of the polymer, which, once reached, can not be extended any further. If this is the case, it would also suggest that termination of the polymerization reaction, the release of the synthase from the polymer, and/or reinitiation of polymerization by the newly released synthase are relatively slow events since no evidence of these reactions are seen within the timecourse of these studies. The phenomenon observed in Figure 15 is not the result of

3 4

decay of the enzyme over the course of the assay since virtually identical results are obtained following a 10 minute preincubation of the synthase at 25°C.

It must also be noted that comparisons of the direct spectrophotometric assays used here and the more common assay involving the use of Ellman's reagent, DTNB, (Ellman, *supra*) in the formation of thiolate of coenzyme-A showed that the values determined by the direct method were approximately 70% of the values determined using Ellman's reagent. This may be due to phase separation occurring in the cuvettes as the relatively insoluble polymer is formed. In support of this notion, a faint haze or opalescence in the cuvette developed during the course of the reaction, particularly at higher substrate concentrations.

PHA synthase purified from insect cells appears to be relatively stable. Examination of activity following storage, in liquid N<sub>2</sub> and at -20°C in the presence of 50% glycerol showed that approximately 50% of synthase activity remained after 7 weeks when stored in liquid N<sub>2</sub> and approximately 75% of synthase activity remained after 7 weeks when stored at -20°C in the presence of 50% glycerol.

The expression of PHA synthase from *A. eutrophus* in a baculovirus expression system results in the synthase constituting approximately 50% of total protein 60 hours post-infection; however, approximately 50-75% of the synthase is observed in the membrane-associated fraction. This elevated level of expression allowed purification of the soluble PHA synthase using a single chromatographic step on HA. The purity of this preparation is estimated to be approximately 90% (intact PHA synthase and 3 proteolysis products).

The initial specific activity of 12 U/mg was approximately 20-fold higher than the most successful previous

efforts at overexpression of *A. eutrophus* PHA synthase. The synthase reported here was isolated from a 250 ml culture with 70% recovery which represents an improvement of 500-fold (1000 U / 64 U x 8 L / 0.25 L) when compared to an 8 L *E. coli* culture with 40% recovery.

5 This high expression level should provide sufficient PHA synthase for extensive structural, functional, and mechanistic studies. Furthermore, it is clear that the baculovirus expression system is an attractive option for isolation of other PHA synthases from various

10 PHA synthase produced in the baculovirus system was of sufficient potency to allow direct spectrophotometric analysis of the hydrolysis of the thioester bond of HBCoA at 232 nm. These assays revealed a lag period of approximately 60 seconds, the length of which was variable and inversely related to enzyme concentration. Such a

15 lag period presumably reflects a slow step in the reaction, perhaps correlating to dimerization of the enzyme, the priming, and/or initiation steps in formation of PHB. Size exclusion chromatographic examination of the PHB synthase native MW indicated two forms of the synthase. One form showed a MW of approximately 100-160 kDa and the other showed a MW of approximately 50-80 kDa; these two

20 forms likely represent the dimer and monomer of PHA synthase, respectively. Similar results have been reported previously in which two forms of approximately 60 and 130 kDa were observed. Comparisons of the direct assay reported here and the indirect assay using DTNB revealed that the former resulted in values that were

25 70% of the values determined by the DTNB indirect assay. Although the reason for this difference has not been examined in detail, it is probable that the apparent phase separation that occurred upon PHB formation in the short pathlength cuvettes used, particularly with high [HBCoA], results in this discrepancy.

30

Enzymatic analyses of the PHA synthase have found that the enzyme has a broad pH optimum centered at pH 8.5; however, the studies described herein have been performed at pH 7.2 to provide comparative values with the results of others. Moreover, the specific

5 activity of this enzyme is dependent upon enzyme concentration which confirms and extends earlier results (Gerngross et al., *supra*).

In studies intended to examine the dependence of activity upon enzyme concentration, it became apparent that the extent of the polymerization reaction is dependent on the amount of enzyme included in the reaction mixture. Specifically, decreasing the amount of enzyme leads not only to decreased velocity of reaction but also to a decreased extent of condensation (Figure 15). One possible explanation is that the enzyme is thermally labile; however, identical

10 assays in which the enzyme is preincubated at 25°C for 10 minutes prior to initiation of the reaction had similar results. Another possibility is that a terminal-length of the polymer is reached precluding further condensations until the particular synthase molecule is released from the terminal-length polymer.

This work clearly demonstrates the value of the baculovirus expression system for the production of *A. eutrophus* PHA synthase and for the potential application to studies of other PHA synthases. Furthermore, the high level of expression obtained using the baculoviral system should allow convenient analysis for substrate-specificity and structure-function studies of PHA synthases

25 from relatively crude insect cell extracts.

## Example 2

Co-expression of rat FAS dehydrase mutant cDNA and PHB synthase gene in insect cells.

30 Expression of a rat FAS DH cDNA in *Sf9* cells has been reported previously (Rangan et al., *J. Biol. Chem.*, 266, 19180 (1991));

37

Joshi et al., *Biochem. J.* **226**, 143 (1993). Once activity of the *phbC* gene product had been established in insect cells (see Example 1), baculovirus clones containing the rat FAS DH- cDNA and BacPAAK6:*phbC* were employed in a double infection strategy to determine if PHB would be produced in insect cells. It was not known if an intracellular pool of R(-)-3-hydroxybutyrate would be stable or available as a substrate for the PHB synthase. In order for the R(-)-3-hydroxybutyryl-CoA to be available as a substrate, the R(-)-3-hydroxybutyryl-CoA released from rat FAS DH- protein must be trapped by the PHB synthase and incorporated into a polymer at a rate faster than  $\beta$ -oxidation, which would regenerate acetyl-CoA. It was also not known if the stereochemical configuration of the 3-hydroxyl group, which must be in the R form, would be recognized as a substrate by PHB synthase. Fortunately, previous biochemical studies on eukaryotic FASs indicated that the R form of 3-hydroxybutyryl-CoA would be generated (Wakil et al., *J. Biol. Chem.* **237**, 687 (1962)).

SDS-PAGE of protein samples from a time course of uninfected, single-infected, and dual-infected S/21 cells was performed (Figure 16). From these data, it is clear that the rat FAS DH mutant and PHB synthase polypeptides are efficiently co-expressed in S/21 cells. However, co-expression results in ~50% reduced levels of both polypeptides compared to S/21 cells that are producing the individual proteins. Western analysis using anti-rat FAS (Rangan et al., *supra*) and anti-PHA synthase antibodies confirmed simultaneous production of the corresponding proteins.

To provide further evidence that PHB was being synthesized in insect cells, *T. ni* cells which had been infected with a baculovirus vector encoding rat FAS DH<sup>0</sup> and/or a baculovirus vector encoding PHA synthase were analyzed for the presence of granules.

38

Infected cells were fixed in paraformaldehyde and incubated with anti-PHA synthase antibodies (Williams et al., *Protein Exp. Purif.* **7**, 203 (1996)). Granules were observed only in doubly infected cells (Williams et al., *App. Environ. Micro.* **62**, 2540 (1996)).

Characterization of PHB production in insect cells. In order to determine if *de novo* synthesis of PHB was occurring in S/21 cells that co-express the rat FAS DH mutant and PHB synthase, fractions of these samples were extracted, the extract subjected to propanolysis, and analyzed for the presence of propylhydroxybutyrate by gas chromatography (Figure 17). A unique peak with a retention time that coincided with a propylhydroxybutyrate standard was detected only in the double infection samples at 48 and 72 hours, in contrast to the individually expressed gene products and uninfected controls, which were negative. These samples were analyzed further by GC/MS to confirm the identity of the product. Figure 18 shows mass spectroscopy data corresponding to the material obtained from peak 10.1 in the gas chromatograph compared to an propylhydroxybutyrate standard. The results show that PHB synthesis is occurring only in S/21 cells co-expressing the rat FAS DH mutant cDNA and the *phbC* gene from *A. eutrophus*. Integration of the peak in the gas chromatograph corresponding to propylhydroxybutyrate revealed that approximately 1 mg of PHB was isolated from 1 liter culture of S/21 cells (approximately 600 mg dry cell weight of S/21 cells). Thus, the ratFAS206 protein effectively replaces the  $\beta$ -ketothiolase and acetoacetyl-CoA reductase functions, resulting in the production of PHB by a novel pathway.

The approach described here provides a new strategy to combine metabolic pathways that are normally engaged in primary anabolic functions for production of polyesters. The premature termination of the normal fatty acid biosynthetic pathway to provide

39

suitably modified acyl-CoA monomers for use in PHA synthesis can be applied to both prokaryotic and eukaryotic expression since the formation of polymer will not be dependent on specialized feedstocks. Thus, once a recombinant PHA monomer synthase is introduced into a prokaryotic or eukaryotic system, and co-expressed with the appropriate PHA synthase, novel biopolymer formation can occur.

### Example 3

#### Cloning and Sequencing of the *vep* ORF1 PKS Gene Cluster

The entire PKS cluster from *Streptomyces venezuelae* was cloned using a heterologous hybridization strategy. A 1.2 kb DNA fragment that hybridized strongly to a DNA encoding an *eryA* PKS  $\beta$ -ketoacyl synthase domain was cloned and used to generate a plasmid for gene disruption. This method generated a mutant strain blocked in the synthesis of the antibiotic. A *S. venezuelae* genomic DNA library was generated, and used to clone a cosmid containing the complete methymycin aglycone PKS DNA. Fine-mapping analysis was performed to identify the order and sequence of catalytic domains along the multifunctional PKS (Figure 19). DNA sequence analysis of the *vep* ORF1 showed that the order of catalytic domains is KSQ/AT/ACP/KS/AT/KR/ACP/KS/AT/DH/KR/ACP. The complete DNA sequence, and corresponding amino acid sequence, of the *vep* ORF1 is shown in Figure 23 (SEQ ID NO:1 and SEQ ID NO:2, respectively).

The sequence data indicated that the PKS gene cluster encodes a polyene of twelve carbons. The *vep* gene cluster contains 5 polypeptide synthase modules, with a loading module at its 5' end and an ending domain at its 3' end. Each of the sequenced modules includes a keto-ACP (KS), an acyltransferase (AT), a dehydratase (DH),

40

a keto-reductase (KR), and an acyl carrier protein domain. The six acyltransferase domains in the cluster are responsible for the incorporation of six acetyl-CoA moieties into the product. The loading module contains a KSQ, an AT and an ACP domain. KSQ refers to a domain that is homologous to a KS domain except that the active site cysteine (C) is replaced by glutamine (Q). There is no counterpart to the KSQ domain in the PKS clusters which have been previously characterized.

The ending domain (ED) is an enzyme which is responsible for the attachment of the nascent polyketide chain onto another molecule. The amino acid sequence of ED resembles an enzyme, HetM, which is involved in *Anabaena* heterocyst formation. The homology between *vep* and HetM suggests that the polypeptide encoded by the *vep* gene cluster may synthesize a polyene-containing composition which is present in the spore coat or cell wall of its natural host, *S. venezuelae*.

### Example 4

To provide a recombinant monomer synthase that generates a saturated  $\beta$ -hydroxyhexanoyl-CoA or unsaturated  $\beta$ -hydroxyhexanoyl-CoA monomer, the linear correspondence between the genetic organization of the Type I macrolide PKS and the catalytic domain organization in the multifunctional proteins is assessed (Donadio et al., *supra*, 1991; Katz et al., *Ann. Rev. Microbiol.* 47, 875 (1993)). First, a DNA encoding a TE is added to the 3' end of an ORF1 of a Type I PKS, preferably the *met* ORF 1 (Figure 6) as recently described by Cortes et al. (*Science*, 268, 1487 (1995) in the erythromycin system. To ensure that the DNA encoding the TE is completely active, DNA encoding a linker region separating a normal ACP-TE region in a PKS, for example the one found in *met* PKS ORF5 (Figure

4 1

5), will be incorporated into the DNA. The resulting vector can be introduced into a host cell and the TE activity, rate of release of the CoA product, and identity of the fatty acid chain determined.

The acyl chain that is most likely to be released is the CoA

5 ester, specifically the 3-hydroxy-4-methyl heptenoylCoA ester, since the fully elongated chain is presumably released in this form prior to macroide cyclization. If the CoA form of the acyl chain is not observed, then a gene encoding a CoA ligase will be cloned and co-expressed in the host cell to catalyze formation of the desired intermediate.

There is clear precedent for release of the predicted premature termination products from mutant strains of macroide-producing *Streptomyces* that produce intermediates in macroide synthesis (Hiber et al., Antimicrob. Agents. Chemother., **34**, 1535 (1990); Kinoshita et al., J. Chem. Soc. Chem. Comm., **14**, 943 (1988)). The structure of these intermediates is consistent with the linear organization of functional domains in macroide PKSs, particularly those related to *eryA*, *tyl*, and *met*. Other known PKS gene clusters include, but are not limited to, the gene cluster encoding 6-methylsalicylic acid synthase (Beck et al., Eur. J. Biochem., **192**, 487 (1990)), soraphen A (Schupp et al., J. Bacteriol., **177**, 3673 (1995), and stigmamycin (Yu et al., J. Bacteriol., **177**, 4792 (1995)).

Once the release of the 3-hydroxy-4-methyl heptenoylCoA ester is established, DNA encoding the extender unit AT in *met* module 1 is replaced to change the specificity from methylmalonylCoA to malonylCoA (Figures 4-6). This change eliminates methyl group branching in the  $\beta$ -hydroxy acyl chain. While comparison of known AT amino acid sequences shows high overall amino acid sequence conservation, distinct regions are readily apparent where significant deletions or insertions have occurred. For

4 2

example, comparison of malonyl and methylmalonyl amino acid sequences reveals a 37 amino acid deletion in the central region of the malonyltransferase. Thus, to change the specificity of the methylmalonyl transferase to malonyl transferase, the *met* ORF1 DNA encoding the 37 amino acid sequence of MMT will be deleted, and the resulting gene will be tested in a host cell for production of the desmethyl species, 3-hydroxyheptenoylCoA. Alternatively, the DNA encoding the entire MMT can be replaced with a DNA encoding an intact MT to affect the desired chain construction.

After replacing MMT with MT, DNA encoding DH/ER will be introduced into DNA encoding *met* ORF1 module 1. This modification results in a multifunctional protein that generates a methylene group at C-3 of the acyl chain (Figure 6). The DNA encoding DH/ER will be PCR amplified from the available *eryA* or *tyl* PKS sequences, including the DNA encoding the required linker regions, employing a primer pair to conserved sequences 5' and 3' of the DNA encoding DH/ER. The PCR fragment will then be cloned into the *met* ORF1. The result is a DNA encoding a multifunctional protein (MT\* DH/ER\*TE\*). This protein possesses the full complement of keto group processing steps and results in the production of heptenoylCoA.

The DNA encoding dehydrase in *met* module 2 is then inactivated, using site-directed mutagenesis in a scheme similar to that used to generate the rat FAS DH described above (Joshi et al., J. Biol. Chem., **268**, 22508 (1993)). This preserves the required (R)-3-hydroxy group which serves as the substrate for PHA synthases and results in a (R)-3-hydroxyheptanoylCoA species.

The final domain replacement will involve the DNA encoding the starter unit acyltransferase in *met* module 1 (Figure 5), to change the specificity from propionyl CoA to acetyl CoA. This shortens the (R)-3-hydroxy acyl chain from heptanoyl to hexanoyl.

The DNA encoding the catalytic domain will need to be generated based on a FAS or 6-methylsalicylic acid synthase model (Beck et al., Eur. J. Biochem. 192, 487 (1990)) or by using site-directed mutagenesis to alter the specificity of the resident *met* PKS propionyltransferase sequence. Limiting the initiator species to acetyl/CoA can result in the use of this starter unit by the monomer synthase. Previous work with macrolide synthases have shown that some are able to accept a wide range of starter unit carboxylic acids. This is particularly well documented for avermectin synthase, where over 60 new compounds have been produced by altering the starter unit substrate in precursor feeding studies (Dutton et al., J. Antibiotics, 44, 357 (1991)).

#### Example 5

To provide a recombinant monomer synthase that synthesizes 3-hydroxyl-4-hexenoic acid, a precursor for polyhydroxyhexenoate, the DNA segment encoding the loading and the first module of the *vep* gene cluster was linked to the DNA segment encoding module 7 of the *tyl* gene cluster so as to yield a recombinant DNA molecule encoding a fusion polypeptide which has no amino acid differences relative to the corresponding amino acid sequence of the parent modules. The fusion polypeptide catalyzes the synthesis of 3-hydroxyl-4-hexenoic acid. The recombinant DNA molecule was introduced into SCP2, a *Streptomyces* vector, under the control of the *act* promoter (pDH5502, Figure 20). A polyhydroxyalkanoate polymerase gene, *phaC1* from *Pseudomonas oleovorans*, was then introduced downstream of the recombinant PKS cluster (pDH5505; Figures 22 and 23). The DNA segment encoding the polyhydroxyalkanoate polymerase is linked to the DNA segment encoding the recombinant PKS synthase so as to yield a fusion polypeptide which synthesizes polyhydroxyhexenoate in *Streptomyces*. Polyhydroxyhexenoate, a biodegradable thermoplastic,

is not naturally synthesized in *Streptomyces*, or as a major product in any other organism. Moreover, the unsaturated double bond in the side chain of polyhydroxyhexenoate may result in a polymer which has superior physical properties as a biodegradable thermoplastic over the known polyhydroxyalkanoates.

The complete disclosure of all patents, patent documents and publications cited herein are incorporated herein by reference as if individually incorporated. The foregoing detailed description and examples have been given for clarity of understanding only. No unnecessary limitations are to be understood therefrom. The invention is not limited to the exact details shown and described for variations obvious to one skilled in the art will be included within the invention defined by the claims.

4 5

**WHAT IS CLAIMED IS:**

1. A baculovirus expression cassette comprising a nucleic acid molecule encoding a polyhydroxyalkanoate synthase operably linked to a promoter functional in an insect cell.
2. The expression cassette of claim 1 wherein the source of the nucleic acid molecule is a bacterium.
3. The expression cassette of claim 2 wherein the bacterium is *Alcaligenes eutrophus*.
4. An expression cassette comprising a nucleic acid molecule encoding a polyhydroxyalkanoate monomer synthase operably linked to a promoter functional in a host cell, wherein the nucleic acid molecule comprises a first DNA segment encoding a first module and a second DNA segment encoding a second module, wherein the DNA segments together encode a polyhydroxyalkanoate monomer synthase, and wherein no more than one DNA segment is derived from the *eryA* gene cluster of *Saccharopolyspora erythraea*.
5. The expression cassette of claim 4 wherein the source of at least one DNA segment is bacterial DNA.
6. A method of providing a polyhydroxyalkanoate synthase, comprising:
  - (a) introducing an expression cassette comprising a DNA molecule encoding a polyhydroxyalkanoate synthase operably linked to a promoter functional in a eukaryotic cell into the eukaryotic cell, wherein the eukaryotic cell is not of plant origin; and
  - (b) expressing the DNA molecule encoding the polyhydroxyalkanoate synthase in the eukaryotic cell.
7. The method of claim 6 wherein the polyhydroxyalkanoate synthase is polyhydroxybutyrate synthase.
8. The method of claim 6 wherein the polyhydroxyalkanoate synthase is derived from a bacterium.
9. The method of claim 8 wherein the bacterium is *Alcaligenes eutrophus*.
10. The method of claim 6 wherein the eukaryotic cell is of insect origin.
11. The method of claim 10 wherein the expression cassette is a baculovirus expression cassette.
12. The method of claim 6 further comprising isolating polyhydroxyalkanoate synthase from the eukaryotic cell.
13. A method of providing a polyhydroxyalkanoate polymer, comprising:
  - (a) introducing into a eukaryotic cell (i) a first expression cassette comprising a DNA segment encoding a fatty acid synthase in which the dehydrase activity is inactivated that is operably linked to a promoter functional in the eukaryotic cell, and (ii) a second expression cassette comprising a DNA segment encoding a

4 6



47

polyhydroxyalkanoate synthase operably linked to a promoter functional in the eukaryotic cell; and  
 (b) expressing the DNA segments so as to yield a polyhydroxyalkanoate polymer in the eukaryotic cell.

14. The method of claim 13 wherein the eukaryotic cell is of insect origin.

15. The method of claim 13 wherein the dehydrase activity is inactivated by mutating the catalytic site.

16. The method of claim 13 wherein the fatty acid synthase is a rat fatty acid synthase.

17. The method of claim 13 wherein the polyhydroxyalkanoate synthase is a polyhydroxybutyrate synthase.

18. The method of claim 13 wherein the fatty acid synthase produces a premature termination product.

19. The method of claim 13 wherein the fatty acid synthase catalyzes the synthesis of D(-)-3-hydroxybutyrate in the eukaryotic cell.

20. The method of claim 13 wherein the polyhydroxyalkanoate polymer is polyhydroxybutyrate.

21. The method of claim 13 wherein the first and second expression cassettes are on different DNA molecules.

48

22. An isolated and purified DNA molecule comprising a first DNA segment encoding a first module and a second DNA segment encoding a second module, wherein the DNA segments together encode a recombinant polyhydroxyalkanoate monomer synthase, and wherein no more than one DNA segment is derived from the *eryA* gene cluster of *Saccharopolyspora erythraea*.

23. The isolated DNA molecule of claim 22 wherein the first DNA segment is derived from the *vcp* gene cluster of *Streptomyces venezuelae*.

24. The isolated DNA molecule of claim 22 wherein the second DNA segment is derived from the *tyl* gene cluster of *Streptomyces*.

25. The isolated DNA molecule of claim 22 wherein the second DNA segment comprises a DNA encoding a thioesterase which is located at the 3' end of the second DNA segment.

26. The isolated DNA molecule of claim 25 wherein the second DNA segment comprises a DNA encoding an acyl carrier protein which is located 5' to the DNA encoding the thioesterase.

27. The isolated DNA molecule of claim 26 wherein the second DNA segment comprises a DNA encoding a linker region, wherein the DNA encoding the linker region is located between the DNA encoding the acyl carrier protein and the DNA encoding the thioesterase.

49

28. The isolated DNA molecule of claim 22 wherein the first DNA segment comprises DNA encoding two acyl transferases, wherein the DNA encoding the first acyl transferase is 5' to the DNA encoding the second acyl transferase.
29. The isolated DNA molecule of claim 28 wherein the second acyl transferase adds acyl groups to malonylCoA.
30. The isolated DNA molecule of claim 22 wherein the first DNA segment comprises a DNA encoding a dehydratase.
31. The isolated DNA molecule of claim 22 wherein the first DNA segment comprises a DNA encoding a dehydratase and an enoyl reductase.
32. The isolated DNA molecule of claim 22 wherein the second DNA segment comprises a DNA encoding an inactive dehydratase.
33. The isolated DNA molecule of claim 22 wherein the first DNA segment comprises a DNA encoding an acyl transferase.
34. The isolated DNA molecule of claim 33 wherein the acyl transferase domain binds an acyl CoA substrate.
35. The isolated DNA molecule of claim 22 comprising a first DNA segment encoding a first module and a second DNA segment encoding a second module, wherein the DNA segments together encode a recombinant polyhydroxyalkanoate monomer synthase, and wherein no more than one DNA

50

36. The isolated DNA molecule of claim 35 wherein the first DNA segment encodes the first module from the *vfp* gene cluster and the second DNA segment encodes module 7 from the *tyl* gene cluster.
37. A method of providing a polyhydroxyalkanoate monomer, comprising:
  - (a) introducing into a host cell a DNA molecule comprising a DNA segment encoding a recombinant polyhydroxyalkanoate monomer synthase operably linked to a promoter functional in the host cell, wherein the recombinant polyhydroxyalkanoate monomer synthase comprises a first module and a second module; and
  - (b) expressing the DNA encoding the recombinant polyhydroxyalkanoate monomer synthase in the host cell so as to generate a polyhydroxyalkanoate monomer.
38. A method of providing a polyhydroxyalkanoate polymer, comprising:
  - (a) introducing into a host cell a first DNA molecule comprising a DNA segment encoding a recombinant polyhydroxyalkanoate monomer synthase operably linked to a promoter functional in a host cell, wherein the recombinant polyhydroxyalkanoate monomer synthase comprises a first module and a second module; introducing into the host cell of step (a) a second DNA molecule comprising a DNA segment encoding a
  - (b)

polyhydroxyalkanoate synthase operably linked to a promoter functional in the host cell; and

(c) expressing the DNAs encoding the recombinant polyhydroxyalkanoate monomer synthase and polyhydroxyalkanoate synthase in the host cell so as to generate a polyhydroxyalkanoate polymer.

39. The method of claim 37 or 38 wherein the first DNA segment encodes the first module from the *rep* gene cluster and the second DNA segment encodes module 7 from the *tyl* P gene cluster.
40. An isolated and purified DNA molecule comprising a first DNA segment encoding a fatty acid synthase and a second DNA segment encoding a module of a polyketide synthase.
41. The isolated DNA molecule of claim 40 wherein the second DNA segment encodes a  $\beta$ -ketoacyl synthase amino-terminal to an acyltransferase which is amino-terminal to a ketoreductase which is amino-terminal to an acyl carrier protein which is amino-terminal to a thioesterase.
42. The isolated DNA molecule of claim 40 wherein the second DNA segment is 3' to the DNA encoding the fatty acid synthase.
43. The isolated DNA molecule of claim 40 wherein the second DNA segment is separated from the first DNA segment by a DNA encoding a linker region.

44. The isolated DNA molecule of claim 41 wherein the DNA encoding the linker region is selected from the group consisting of *tyl* ORF1 ACP<sub>1</sub>-KS<sub>2</sub>, *tyl* ORF1 ACP<sub>2</sub>-KS<sub>3</sub>, *tyl* ORF3 ACP<sub>3</sub>-KS<sub>6</sub>, *eryA* ORF1 ACP<sub>1</sub>-KS<sub>1</sub>, *eryA* ORF1 ACP<sub>2</sub>-KS<sub>2</sub>, *eryA* ORF2 ACP<sub>3</sub>-KS<sub>4</sub>, and *eryA* ORF2 ACP<sub>3</sub>-KS<sub>6</sub>.

45. The isolated DNA molecule of claim 40 comprising a first DNA segment encoding a fatty acid synthase and a second DNA segment encoding a module of a polyketide synthase.

46. A method of providing a polyhydroxyalkanoate monomer, comprising:
  - (a) introducing into a host cell a DNA molecule comprising a first DNA segment encoding a fatty acid synthase and a second DNA segment encoding a polyketide synthase, wherein the first DNA segment is 5' to the second DNA segment, wherein the first DNA segment is operably linked to a promoter functional in the host cell, and wherein the first DNA segment is linked to the second DNA segment so that the linked DNA segments express a fusion protein; and
  - (b) expressing the DNA molecule in the host cell so as to generate a polyhydroxyalkanoate monomer.

47. The method of claim 46 wherein the host cell is selected from the group consisting of insect cells, *Streptomyces* cells and *Pseudomonas* cells.

48. The method of claim 46 wherein the DNA encoding the fatty acid synthase is eukaryotic in origin.

5 3

49. The method of claim 46 wherein the DNA encoding the fatty acid synthase is prokaryotic in origin.
50. The method of claim 46 wherein the DNA encoding the polyketide synthase module is derived from DNA encoding the *tyl* module F.
51. An expression cassette comprising a DNA molecule comprising a DNA segment encoding a fatty acid synthase and a polyhydroxyalkanoate synthase.
52. A method of providing a polyhydroxyalkanoate monomer synthase, comprising:
  - (a) introducing an expression cassette comprising a DNA molecule encoding a polyhydroxyalkanoate synthase operably linked to a promoter functional in a host cell, wherein the DNA comprises a first DNA segment encoding a first module and a second DNA segment encoding a second module wherein the DNA segments together encode a polyhydroxyalkanoate monomer synthase; and
  - (b) expressing the DNA molecule in the host cell.
53. An isolated and purified DNA molecule comprising a DNA segment encoding a *Streptomyces venezuelae* polyketide synthase.
54. The isolated DNA molecule of claim 53 wherein the DNA segment comprises SEQ ID NO:1.

5 4

55. The isolated DNA molecule of claim 53 wherein the DNA segment encodes a polypeptide having an amino acid sequence comprising SEQ ID NO:2.
56. The expression cassette of claim 4 wherein the first DNA segment encodes the first module from the *wcp* gene cluster and the second DNA segment encodes module 7 from the *tyl* P gene cluster.
57. The expression cassette of claim 4 further comprising a third DNA segment encoding a polyhydroxyalkanoate synthase.
58. The method of claim 37 wherein the DNA molecule further comprises a DNA segment encoding a polyhydroxyalkanoate synthase.
59. The isolated DNA molecule of claim 22 or 36 further comprising a DNA segment encoding a polyhydroxyalkanoate synthase.
60. The method of claim 53 wherein the expression cassette further comprises a second DNA molecule encoding a polyhydroxyalkanoate synthase

Figure 1

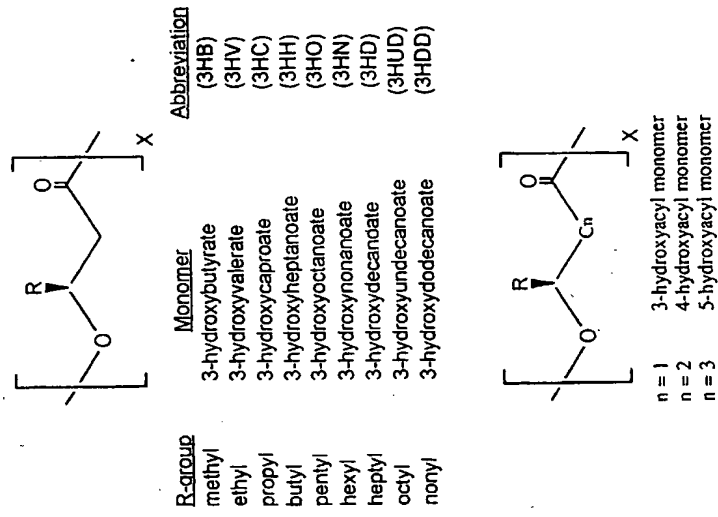
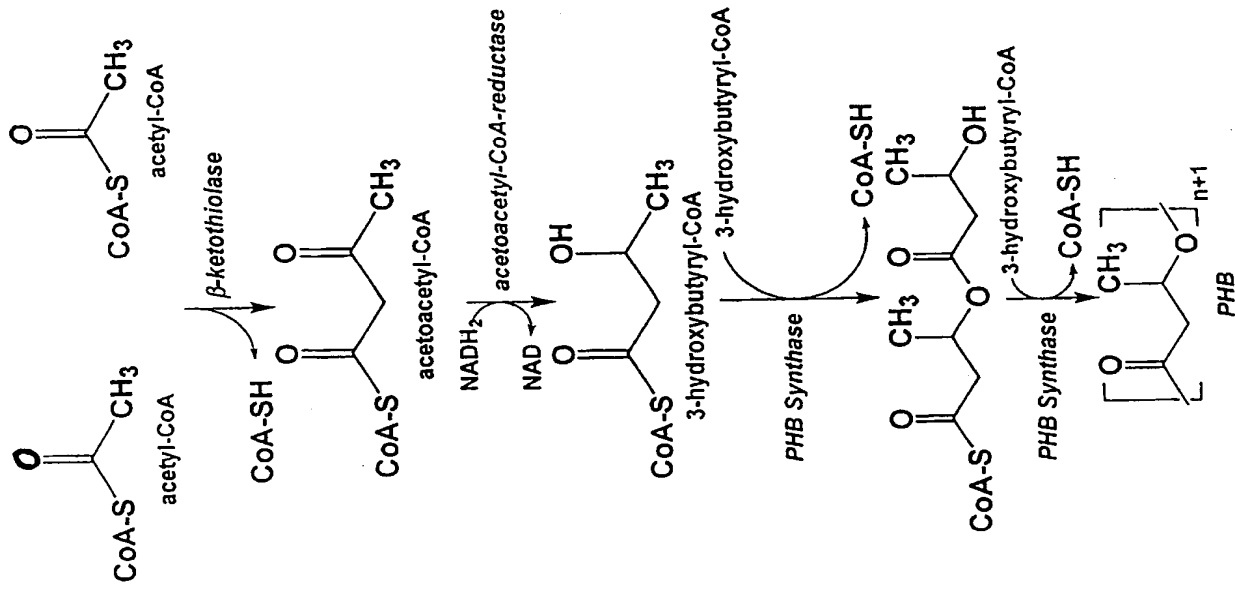


Figure 2

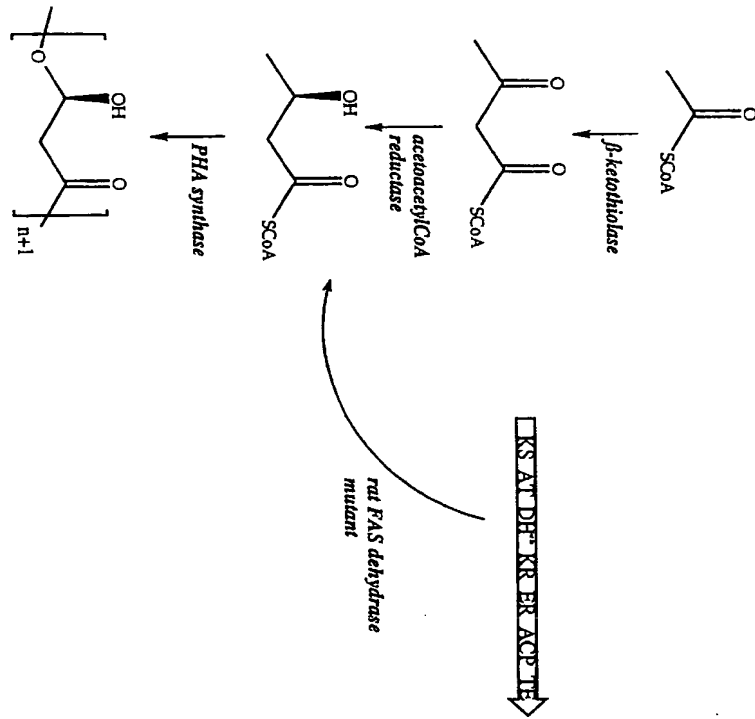


Figure 3

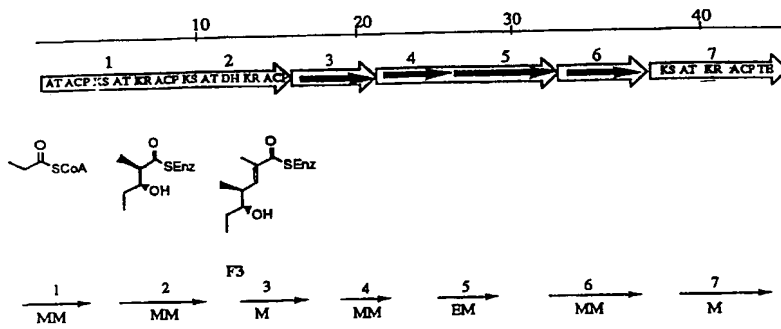
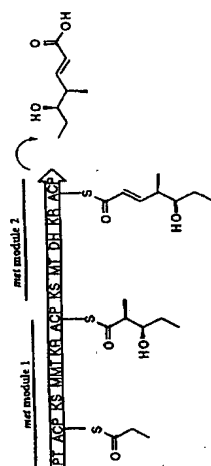


Figure 4



1. introduce TE domain and establish release of acyl CoA ester
2. change MMT to MT domain in module 1
3. introduce DHYER (or DH only) domain into module 1
4. inactivate DH domain in module 2
5. replace PT starter domain with AT in module 1

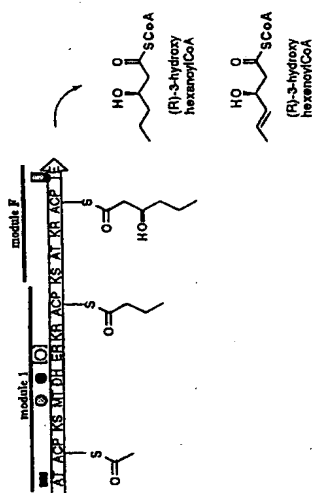


Fig 6

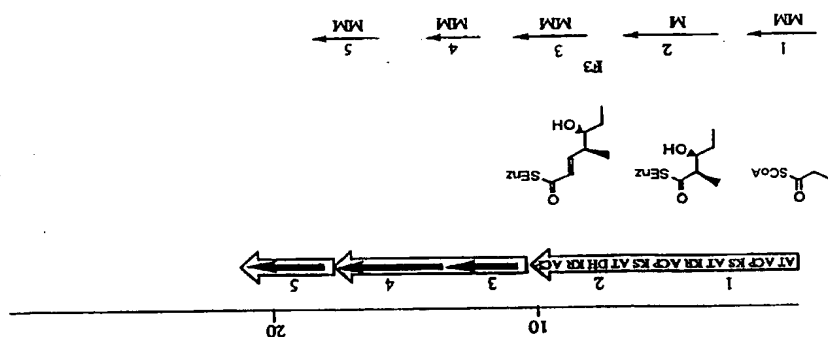


Figure 5

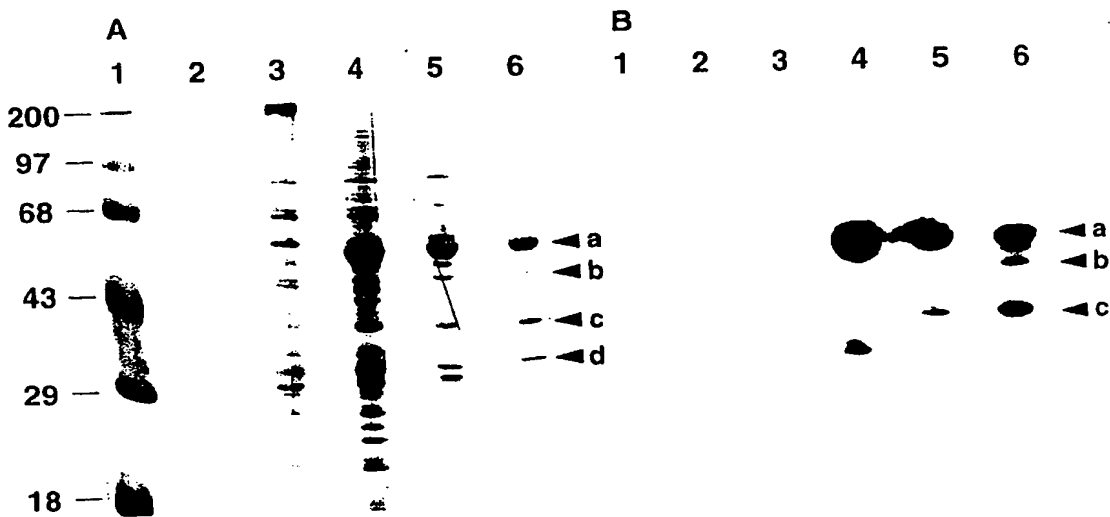


Figure 7

Figure 8

N-terminal sequence determined for PHA synthase				
	1	10	20	25
a	MATGGAASSTQEGKSQPFKVTTPGP-			
b	AAASTQEGKSQPFKVTTPGP-			
c	STQEGKSQPFKVTTPGP-			



Figure 9

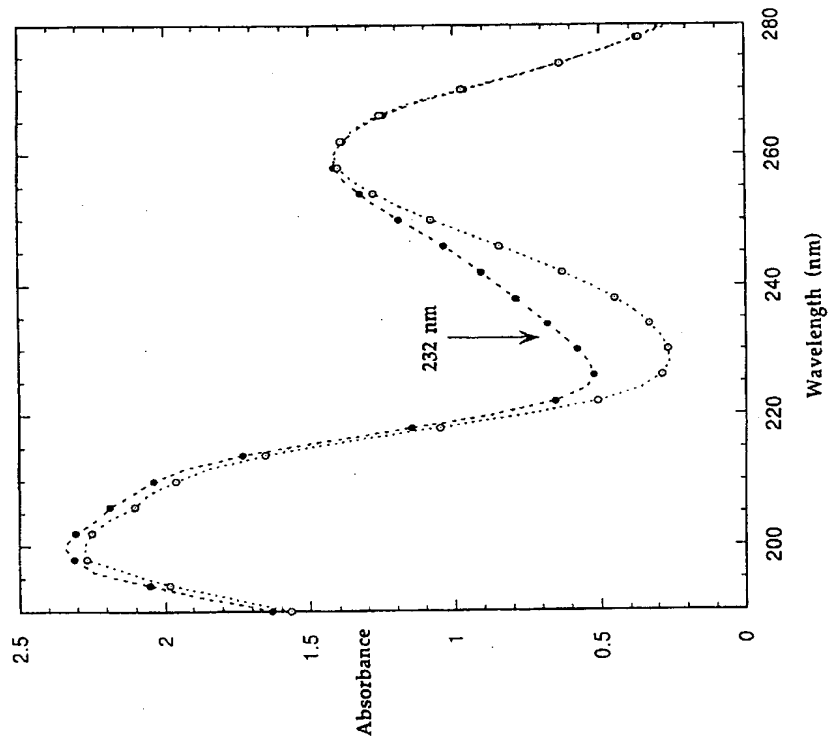


Figure 10

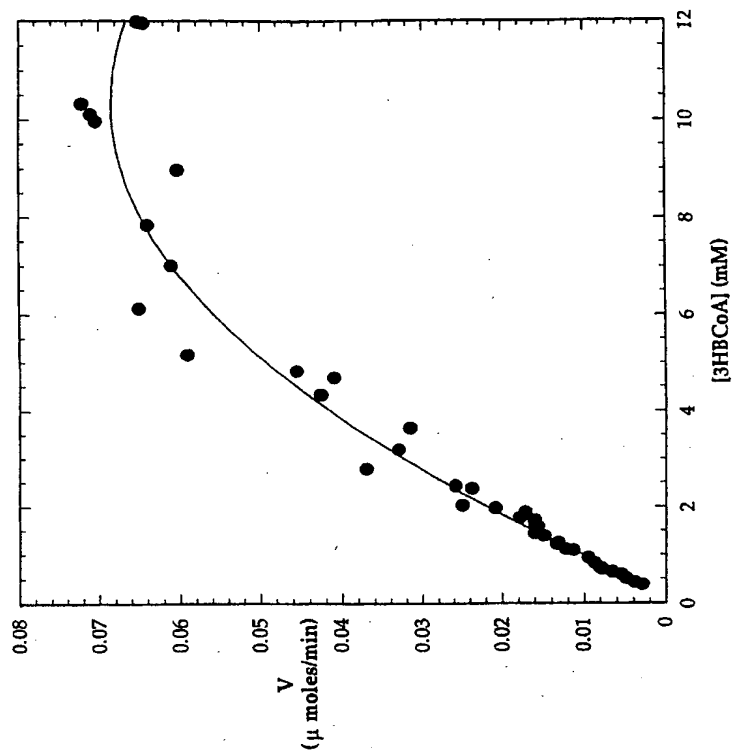


Figure 11

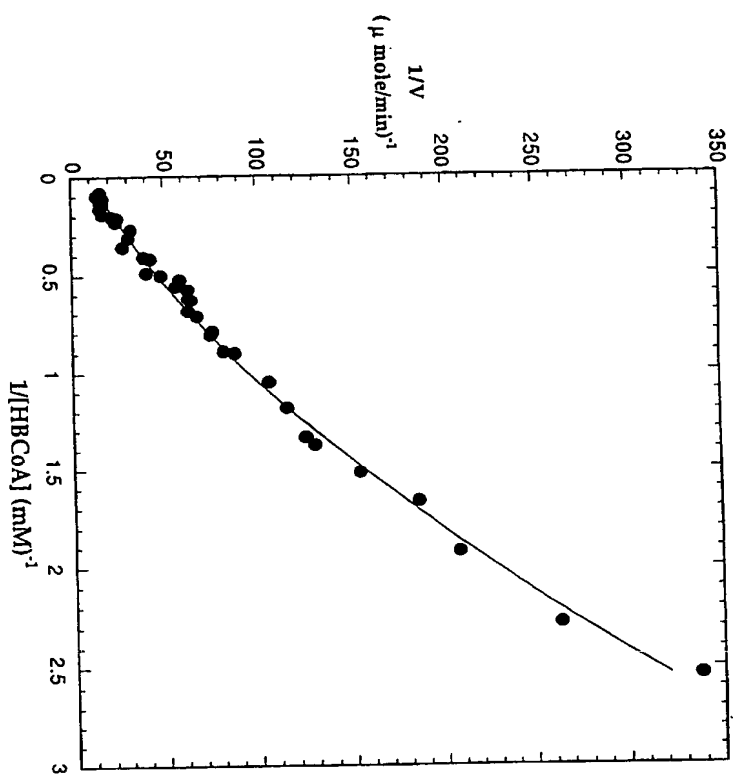


Figure 12

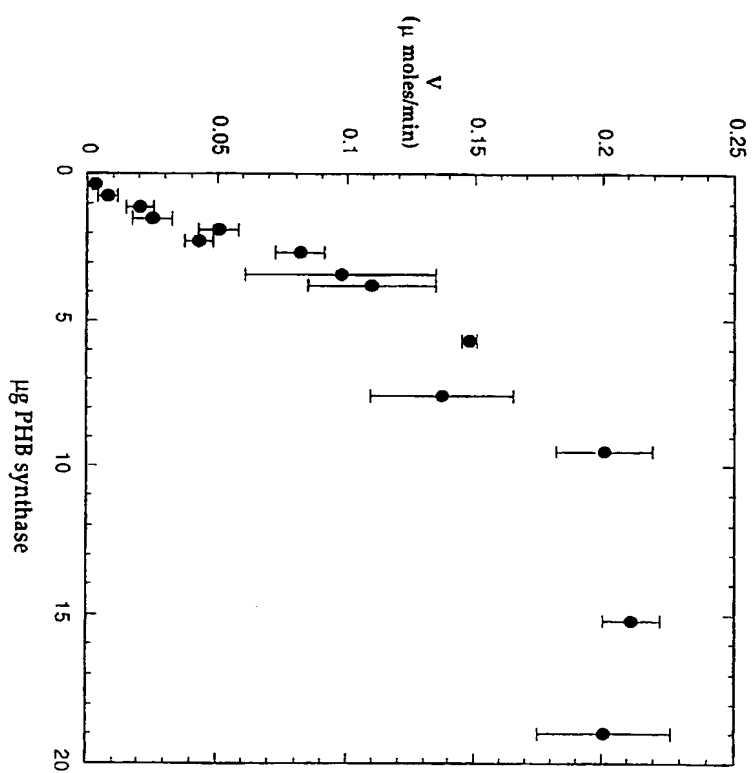


Figure 13

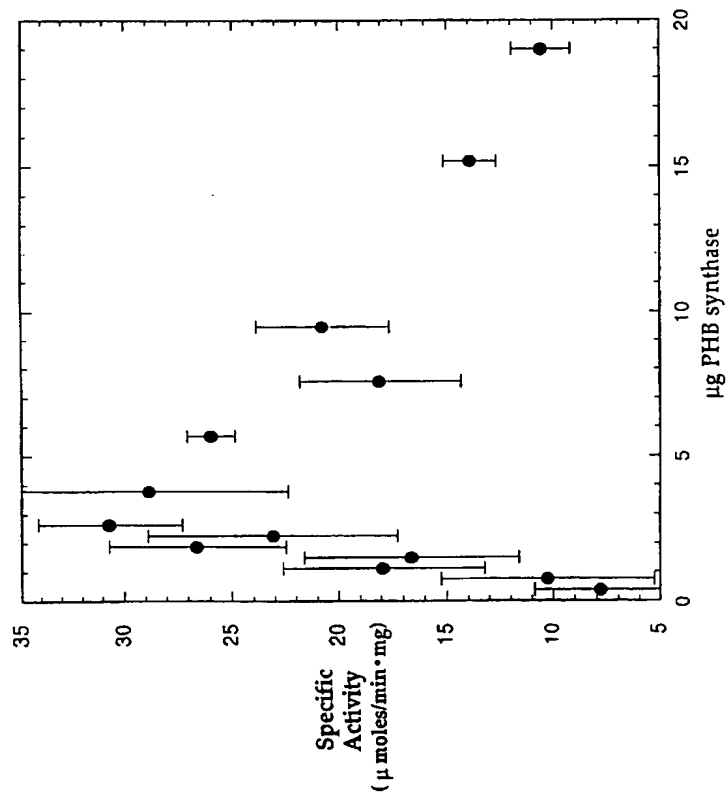


Figure 14

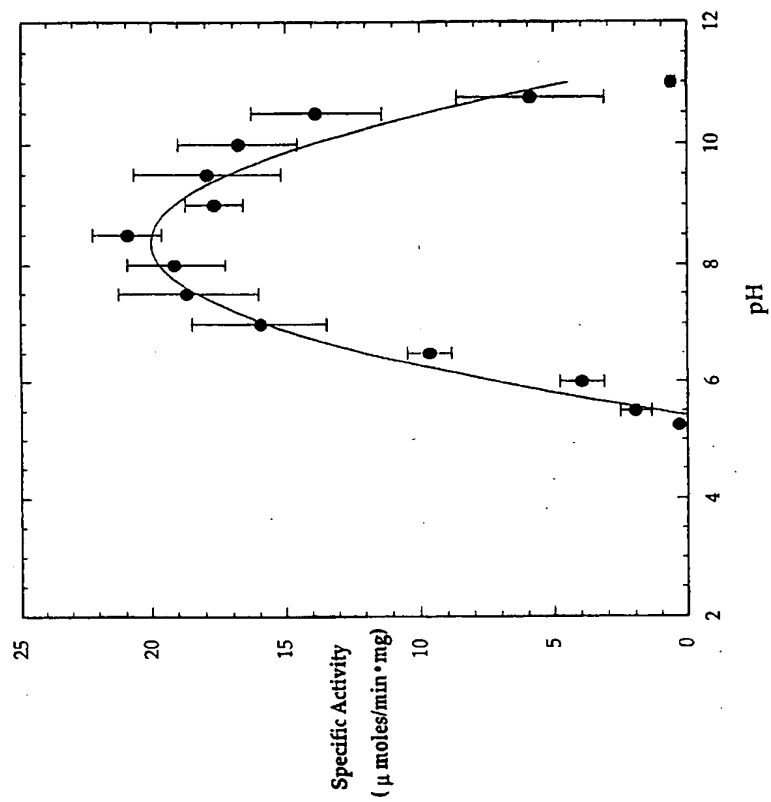


Figure 15

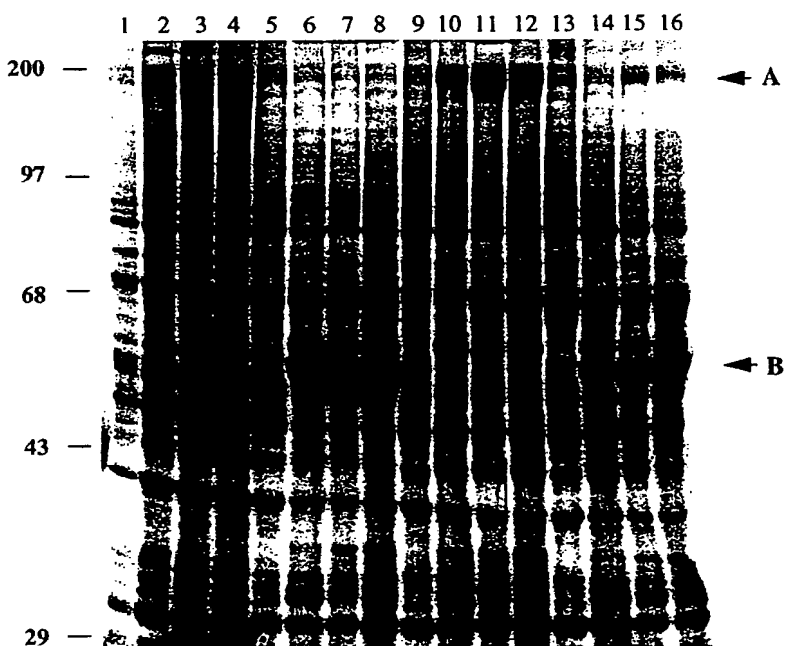
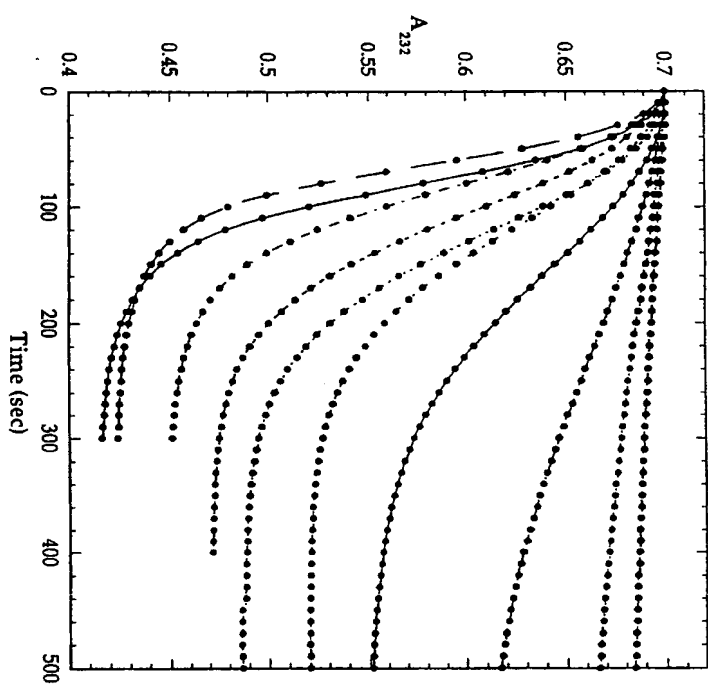


Figure 16

18/33

WO 97/22711

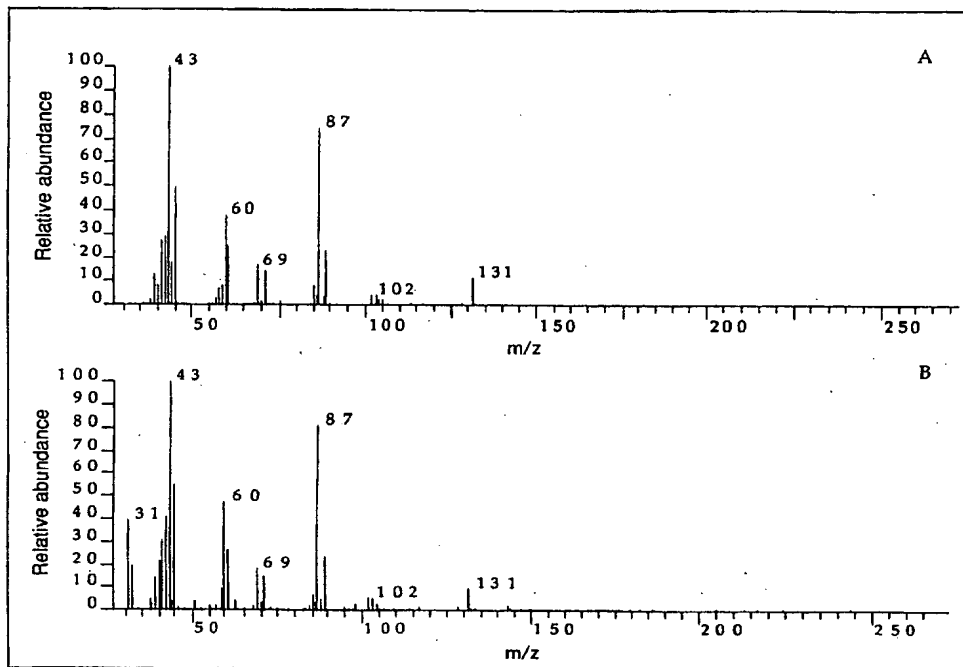


Figure 18

17/33

WO 97/22711

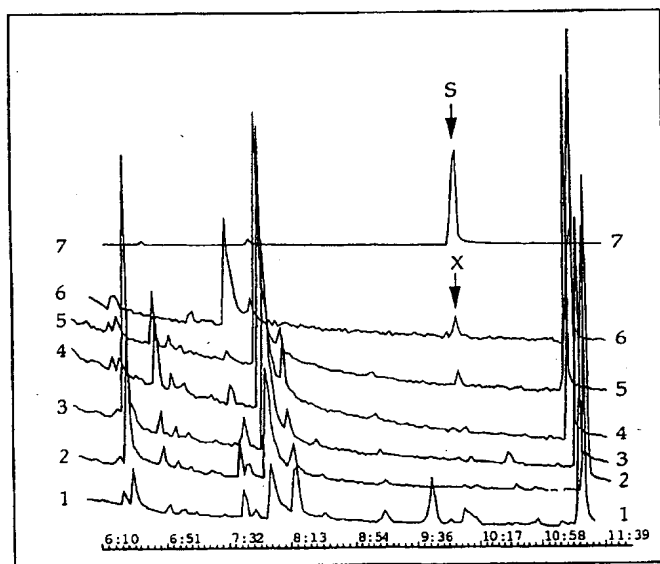


Figure 17

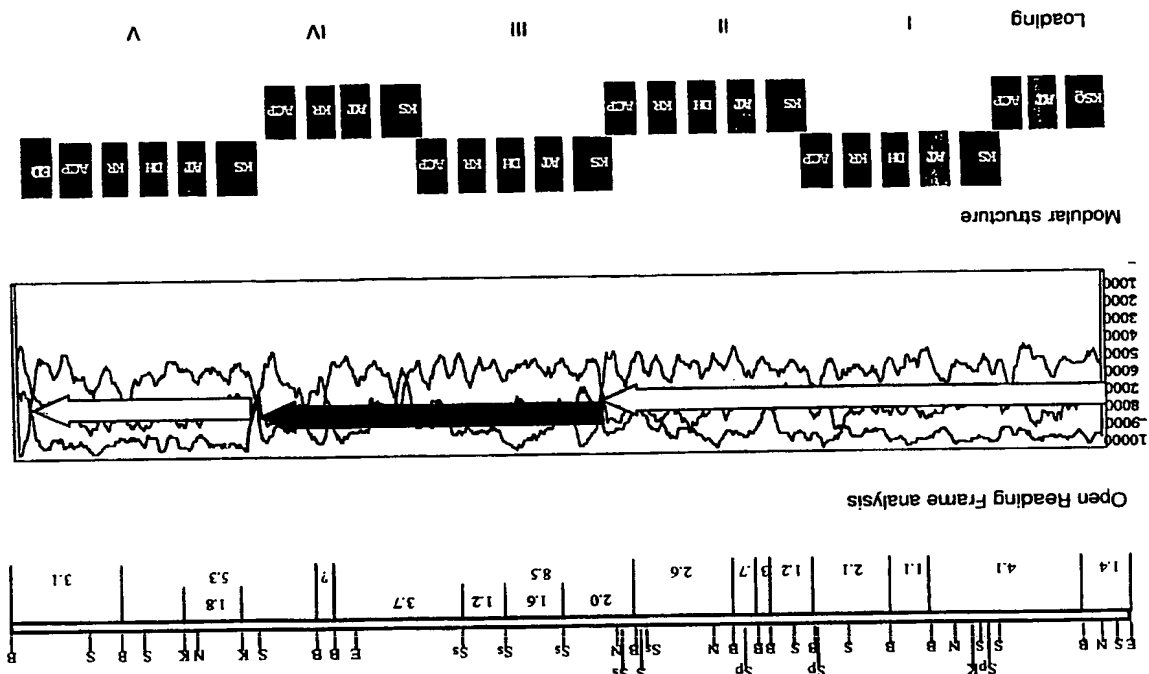


Figure 19

### The Map of *uep* Cluster

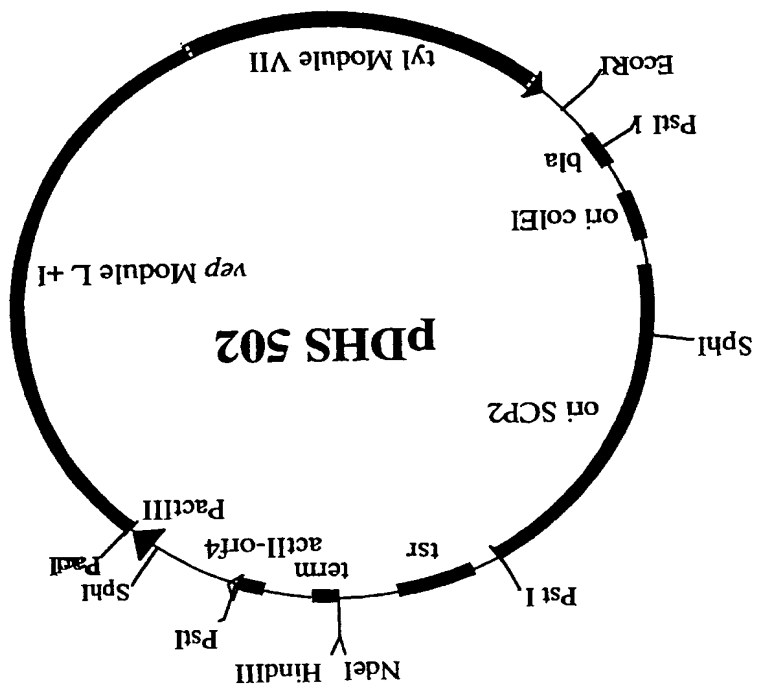
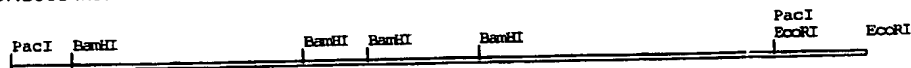


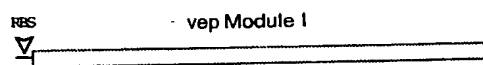
Figure 20

# pDHS505 Construction Procedure

Restriction map of pDHS505 insert

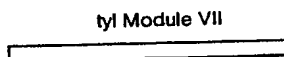


Major steps in the construction

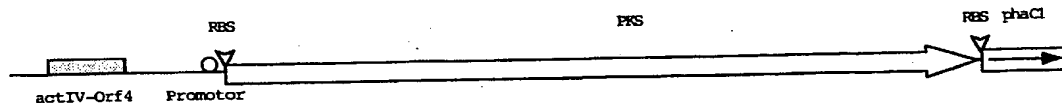
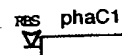


1. A Streptomyces Ribosome Binding Site (RBS) was introduced 6 nucleotides upstream of the translation start site to enhance gene translation in the host.

2. The *tyl* Module VII was recombined by a BamHI site with the vep Module I to give out a complete polyketide synthase Open Reading Frame (ORF) with a Thioesterase at 3'-end.



3. The *phaC1* gene was transcriptionally coupled with the PKS gene. The second Ribosome Binding Site (RBS) was introduced to facilitate the gene translation.



4. The whole expression construct was put under the control of act promotor and the *actII-Orf4* provides an activator which enhances the transcription and expression of the genes.

Figure 22

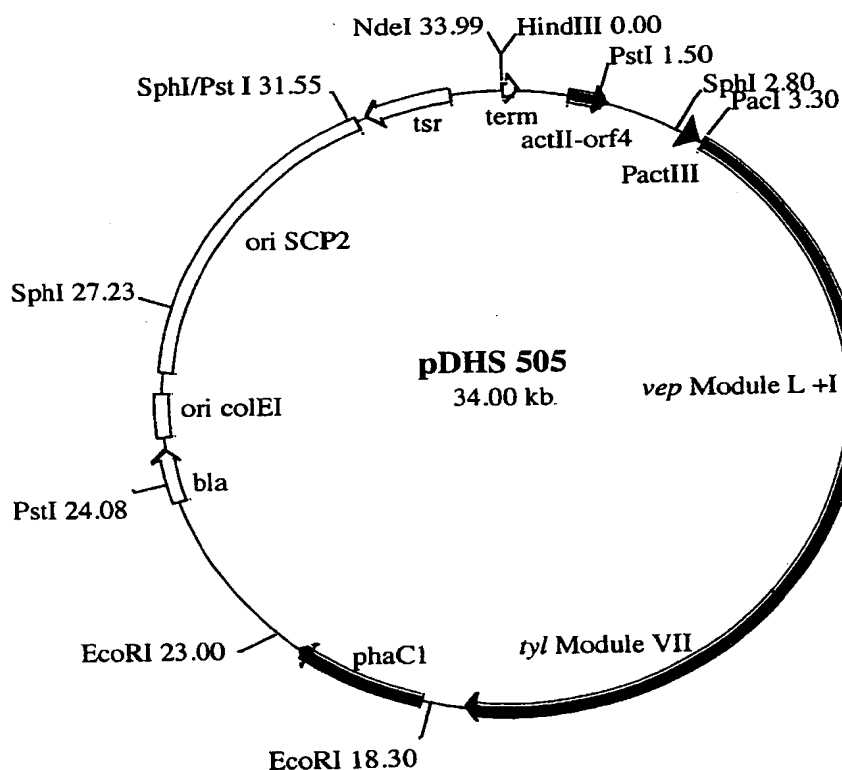


Figure 21

[illegible]

Figure 23

1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479	1480	1481	1482	1483	1484	1485	1486	1487	1488	1489	1490	1491	1492	1493	1494	1495	1496	1497	1498	1499	1500	1501	1502	1503	1504	1505	1506	1507	1508	1509	1510	1511	1512	1513	1514	1515	1516	1517	1518	1519	1520	1521	1522	1523	1524	1525	1526	1527	1528	1529	1530	1531	1532	1533	1534	1535	1536	1537	1538	1539	1540	1541	1542	1543	1544	1545	1546	1547	1548	1549	1550	1551	1552	1553	1554	1555	1556	1557	1558	1559	1560	1561	1562	1563	1564	1565	1566	1567	1568	1569	1570	1571	1572	1573	1574	1575	1576	1577	1578	1579	1580	1581	1582	1583	1584	1585	1586	1587	1588	1589	1590	1591	1592	1593	1594	1595	1596	1597	1598	1599	1600
1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479	1480	1481	1482	1483	1484	1485	1486	1487	1488	1489	1490	1491	1492	1493	1494	1495	1496	1497	1498	1499	1500	1501	1502	1503	1504	1505	1506	1507	1508	1509	1510	1511	1512	1513	1514	1515	1516	1517	1518	1519	1520	1521	1522	1523	1524	1525	1526	1527	1528	1529	1530	1531	1532	1533	1534	1535	1536	1537	1538	1539	1540	1541	1542	1543	1544	1545	1546	1547	1548	1549	1550	1551	1552	1553	1554	1555	1556	1557	1558	1559	1560	1561	1562	1563	1564	1565	1566	1567	1568	1569	1570	1571	1572	1573	1574	1575	1576	1577	1578	1579	1580	1581	1582	1583	1584	1585	1586	1587	1588	1589	1590	1591	1592	1593	1594	1595	1596	1597	1598	1599	1600
1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479	1480	1481	1482	1483	1484	1485	1486	1487	1488	1489	1490	1491	1492	1493	1494	1495	1496	1497	1498	1499	1500	1501	1502	1503	1504	1505	1506	1507	1508	1509	1510	1511	1512	1513	1514	1515	1516	1517	1518	1519	1520	1521	1522	1523	1524	1525	1526	1527	1528	1529	1530	1531	1532	1533	1534	1535	1536	1537	1538	1539	1540	1541	1542	1543	1544	1545	1546	1547	1548	1549	1550	1551	1552	1553	1554	1555	1556	1557	1558	1559	1560	1561	1562	1563	1564	1565	1566	1567	1568	1569	1570	1571	1572	1573	1574	1575	1576	1577	1578	1579	1580	1581	1582	1583	1584	1585	1586</														

Figure 23 cont.



3085 CTC GGC GGC CAC CTG GGC CAC CTG CTC ACC GGC GGC AGC GGC GAG ACC GCA TCG GGC CAC 3084  
996 L A H L G D L L T G S G E T G S A D 1015  
3085 GGC ATA CCG GGC GGC GGC ACC GGC GGC GGC ACC GGC GGC GGC GGC GGC GGC GGC GGC 3124  
1016 G I P A T P A D T T A E P I A I I G M 1035  
3125 CTC TGC CAC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 3184  
1036 A C K Y P G G V T S P E D L W R L V A E 1055  
3185 GGC CAC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 3244  
1056 G R D A V S G L P T D R G W D E D L F D 1075  
3245 GGC CAC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 3304  
1076 A D P R S G K S S V R E G G F L H D A 1095  
3305 CTC TGC TGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 3364  
1096 A L F D A G F P G I S P R E A L G M D P 1115  
3365 CAC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 3424  
1116 Q R L L L E T A W E A V E R A A L D P 1135  
3425 GAA GGC CAC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 3484  
1136 E G L K G S R T A V F V G A T A L D Y G 1155  
3485 CTC TGC TGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 3544  
1156 P R M H D G A E G V E G H L L T G T P 1175  
3545 GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 3604  
1176 S V M S G R I A Y Q L G L T G P A V T V 1195  
3605 GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 3664  
1196 D T A C S S L V A L H L A V R S L R Q 1215  
3665 GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 3724  
1216 C E S S L A L A G A T V H S T P G M F 1235  
3725 GTC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 3784  
1236 V E F S R Q R G L A A D G R S K A F S D 1255  
3785 TCC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 3844  
1256 S A D G T S W A E G V G L L V V E R L S 1275  
3845 CAC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 3904  
1276 D A E R N G H P V L A V I R G S A V N Q 1295  
3905 GAC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 3964  
1296 D G A L A D A G L T A P N G P S Q Q R V I R 1315  
3965 GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 4024  
1316 Q A L A D A G L T P A D V D A V E A H G 1335  
4025 AGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 4084  
1336 T G R L L G D P I E A E A I L G T Y G R 1355  
4085 GAC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 4144  
1356 D R G E A P L Q L G S L K S N I G H A 1375  
4145 GAC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 4204  
1376 Q A A A G V G G L I K M V L A H R H G V 1395  
4205 GTC CAC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 4264  
1396 L P R T L H V D R P T T R V D W E A G G 1415  
4265 GTC CAC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 4324  
1416 V E L L T E E R E W P E T G R P R A A 1435  
4325 ATC TGC TGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 4384  
1436 I S S F G I S G T N A H I V V E Q A P E 1455  
4385 GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 4444  
1456 A G E A A V T T A P E A G E A A 1475  
4445 GAC ACC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 4504  
1476 D T A T T T P A A V G V P E P V R A P 1495  
4505 GTC GTC TGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 4564  
1496 V V A S R R A A A L R A Q A V R L R T 1515

Figure 23 cont.

4565 TTC CTC CAC GGC GGC CAC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 4624  
1516 F L D G R P D V T V A D L G R S L A A R 1535  
4625 ACC CTC TAC GAC ACC GGC CTC ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC 4684  
1536 T A F E H K A L T T A T R D E L L A G 1555  
4685 CTC CAC GGC CTC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 4744  
1556 L D A L G R G E Q A T G L V T G E P A R 1575  
4745 GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 4804  
1576 A G R T A F L F T G Q A Q R V A M G E 1595  
4805 GAA CTC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 4864  
1596 E L R A A H P V F A A A L D T V Y A A L 1615  
4865 CAC CTC CAC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 4924  
1616 D R H L D R P L R E I V A A G E L D L 1635  
4925 ACC GGC TAC ACC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 4984  
1636 T A Y T P A L P A F E V A L F R L L E 1655  
4985 CAC GGC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC 5044  
1656 H H G L V P D L L T G H S V G E I A A A 1675  
5045 CAC GGC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC 5104  
1676 H V A G V L S L D D A A R L V T A R G R 1695  
5105 CTC ATG CAG TCG CTC CAC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 5164  
1696 L H Q S A R E G G A M I A V Q A G A E 1715  
5165 GTC CTC CAC TCC CTC CAC GGC TAC CAC GGC GGC GGC GGC GGC GGC GGC 5224  
1716 V V E S L K G Y E G R V A A A V N C P 1735  
5225 ACC GGC GTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC 5284  
1736 T A V V V S G D A D A A E I R A W A 1755  
5285 GGC CAC GGC CAC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 5344  
1756 G R G R T R R L R V S H A F S P H H 1775  
5345 GAC GGC CTC CTC CAC GGC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC 5404  
1776 D V L D E F L R V A E G L T F E P R 1795  
5405 ATC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC 5464  
1796 I P V T S T V T G A L V T S G E L T S P 1815  
5465 CTC TAC TCG CTC CAC CAC ACC CTC CTC CTC CTC CTC CTC CTC CTC CTC 5524  
1816 A Y W V D Q I R R P V R P L D A V R T L 1835  
5525 GGC CAC CAC GGC ACC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC 5584  
1836 A A Q D A T V L V E I G P D A V L T A 1855  
5585 GGC CAC GGC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC 5644  
1856 A E A L A P G T D A P D A R D V T V 1875  
5645 CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC 5704  
1876 P L L R A G R P E P E T L A A G L A T A 1895  
5705 CAT CTC CAC GGC CAC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC 5764  
1896 H V H G A P L D R A S F P D G R T D 1915  
5765 CTC CAC TAC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC 5824  
1916 L P T I A F R R E H Y W L T P E A R T D 1935  
5825 GGC CAC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC 5884  
1936 A R A L G F D P A R H P L L T T T V E V 1955  
5885 GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 5944  
1956 A G G D G V L L T G R L S L T D Q P W L 1975  
5945 CAC CAC ATG CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC 6004  
1976 A D H M V N G A V L P A T A F L A 1995  
6005 CTC GGC GGC CAC CAC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC CTC 6064  
1996 L A A G D H V G A V R V E L T L E A P 2015  
6065 CTC CTC CTC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC 6124  
2016 L V L P E R C A V R I Q V G S G D G 2035

Figure 23 Cont.

5125	S	P	A	G	R	T	F	G	V	Y	S	T	P	D	S	G	D	T	G	D	P	A	6184		
2036	S	P	A	G	R	T	F	G	V	Y	S	T	P	D	S	G	D	T	G	D	P	A	2055		
5185	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	6244	
2056	D	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	2075
5185	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	6304	
5245	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	6364	
2076	T	E	S	D	H	P	G	T	D	G	D	G	S	A	A	W	P	P	P	A	A	A	2095		
5245	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	6364	
5305	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	6424	
2096	A	T	A	T	P	P	L	D	G	V	I	D	R	L	A	E	L	G	Y	G	Y	L	2115		
5305	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	6484	
6465	P	A	F	Q	L	T	G	L	M	R	D	G	A	D	T	P	L	A	E	A	A	A	2135		
2116	G	P	A	F	Q	L	T	G	L	M	R	D	G	A	D	T	P	L	A	E	A	A	6544		
6485	I	R	L	P	A	Q	H	E	S	A	G	L	F	G	V	H	P	A	L	A	A	A	2175		
2155	L	D	A	A	L	H	P	I	V	L	E	G	N	S	A	A	G	A	C	D	A	A	6604		
6545	I	R	L	P	A	Q	H	E	S	A	G	L	F	G	V	H	P	A	L	A	A	A	2195		
2176	A	D	T	D	A	T	D	R	I	R	L	P	F	A	W	A	G	V	T	L	A	A	6664		
6605	A	D	T	D	A	T	D	R	I	R	L	P	F	A	W	A	G	V	T	L	A	A	2215		
2195	H	A	E	G	A	T	A	L	R	V	R	I	T	P	A	C	T	D	T	V	A	A	6724		
6665	H	A	E	G	A	T	A	L	R	V	R	I	T	P	A	C	T	D	T	V	A	A	2235		
2216	T	L	R	L	L	T	D	T	T	G	A	P	V	A	T	V	E	S	L	T	L	A	6784		
6725	A	P	A	V	A	K	D	R	L	G	T	T	A	G	R	V	D	D	A	L	L	F	6844		
2236	A	P	A	V	A	K	D	R	L	G	T	T	A	G	R	V	D	D	A	L	L	F	6904		
6785	A	P	A	V	A	K	D	R	L	G	T	T	A	G	R	V	D	D	A	L	L	F	6964		
2256	T	V	W	T	E	T	G	T	P	E	P	A	G	R	G	A	V	E	V	A	E	V	2275		
6905	T	V	W	T	E	T	G	T	P	E	P	A	G	R	G	A	V	E	V	A	E	V	2295		
2276	E	L	D	V	D	L	A	G	L	G	D	L	V	E	L	G	A	A	D	V	A	A	6904		
6945	E	L	D	V	D	L	A	G	L	G	D	L	V	E	L	G	A	A	D	V	A	A	6964		
6965	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	7024	
2296	V	L	R	A	D	R	W	T	L	D	G	D	P	S	A	A	A	R	T	A	A	A	2315		
7024	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	7084	
6965	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	7144	
2316	V	R	T	L	A	T	V	Q	E	F	L	S	E	P	R	F	D	G	S	A	A	A	2335		
7025	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	7204	
2336	R	L	V	C	V	T	R	G	A	V	A	L	L	P	D	G	E	D	V	A	T	S	7264		
7085	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	7324	
2356	L	A	T	G	P	L	W	G	D	L	V	R	S	A	Q	S	E	N	P	G	R	7384			
7145	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	7444	
2376	L	F	L	L	D	L	G	E	G	E	R	D	G	A	E	E	E	L	I	A	A	A	7504		
7205	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	7564	
2396	R	A	A	T	A	A	G	D	E	P	Q	L	A	A	R	D	G	R	L	L	A	A	7624		
7265	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	7684	
2416	P	R	L	A	R	T	A	A	L	S	S	E	D	T	A	G	G	A	D	R	A	A	7744		
7285	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	7804	
2416	P	R	L	A	R	T	A	A	L	S	S	E	D	T	A	G	G	A	D	R	A	A	7864		
7325	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	7924	
2436	F	G	P	D	G	T	V	L	V	T	Q	G	G	L	G	A	L	L	L	L	A	A	7984		
7385	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	8044	
2456	A	R	H	L	V	E	R	H	G	G	V	R	R	L	L	L	V	S	R	R	G	A	8104		
7445	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	8164	
2476	A	D	A	P	G	A	A	D	L	G	E	D	L	A	G	L	G	A	A	E	V	A	8224		
7445	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	8284	
2496	A	F	A	A	A	D	A	A	D	R	E	S	L	A	R	A	I	A	T	V	A	A	8344		
7505	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	8404	
7565	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	8464	
2516	P	A	E	H	P	L	T	A	V	V	H	T	A	G	V	V	D	D	A	T	A	A	8524		
7625	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	8584	
7685	A	P	R	E	W	T	R	H	V	S	C	R	D	A	C	M	L	G	E	G	D	P	A	8644	
2536	V	E	A	L	T	P	E	R	L	D	A	V	L	R	P	K	V	D	A	A	A	A	8704		

Figure 23 cont.

7685	W	N	L	H	E	L	T	K	D	L	R	L	D	A	F	V	L	P	S	G	L	7714
7516	V	S	G	I	V	G	T	A	G	Q	A	N	Y	A	A	A	N	T	G	L		7517
7805	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			7804
2396	D	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G		2397
7865	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			7864
7616	L	W	D	G	T	H	G	M	G	T	L	G	A	A	D	L	A	R	W			7615
7923	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			7924
2635	S	R	A	G	I	T	P	L	T	P	L	Q	G	L	A	L	F	D	A	A		2636
7985	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			7984
2656	V	A	R	D	A	L	L	V	P	A	G	L	R	P	T	A	H	R	G			2657
9045	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			9044
2616	T	D	G	Q	P	P	A	L	W	R	G	L	V	R	A	F	R	R	A			2617
8105	D	D	G	Q	P	P	A	L	W	R	G	L	V	R	A	F	R	R	A			8104
2586	A	R	T	A	A	E	A	A	D	T	T	G	W	L	S	G	L	A	A			2587
8165	S	P	E	E	R	R	S	T	A	V	T	L	V	T	G	V	V	A	D			8164
2716	Q	S	P	E	E	R	R	S	T	A	V	T	L	V	T	G	V	V	A	D		2715
8225	V	L	G	H	A	D	S	A	A	V	G	A	E	R	S	F	K	D	L	G		8224
2736	V	L	G	H	A	D	S	A	A	V	G	A	E	R	S	F	K	D	L	G		2735
8285	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8284
2756	F	D	S	L	A	G	V	E	L	R	N	R	L	N	A	A	T	G	L	R		2755
8345	L	P	A	T	T	V	F	D	H	P	S	P	A	A	L	A	L	A	S	H	L	8344
2776	L	P	A	T	T	V	F	D	H	P	S	P	A	A	L	A	L	A	S	H	L	2775
8405	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8404
2796	A	Q	V	P	R	C	M	L	R	E	G	T	A	A	T	V	V	A	E	R		2795
8465	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8464
8525	A	S	F	G	D	R	A	T	D	D	D	D	P	I	A	I	V	G	M	A		8524
2816	G	A	S	F	G	D	R	A	T	D	D	D	P	I	A	I	V	G	M	A		2815
8525	A	S	F	G	D	R	A	T	D	D	D	D	P	I	A	I	V	G	M	A		8524
2836	C	R	Y	P	G	V	S	S	P	E	D	L	W	R	L	V	A	E	G			2835
8585	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8584
2856	T	D	A	I	S	E	F	P	V	N	R	G	W	D	L	E	S	L	Y	D		8583
8645	S	P	E	E	R	R	S	T	A	V	T	L	V	T	G	V	V	A	D			8644
2876	P	D	F	E	S	K	G	T	T	Y	C	R	E	G	G	F	L	E	G	A		8643
8705	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8704
2886	G	D	F	D	A	A	F	F	G	I	S	P	R	E	A	L	V	H	D	P		2885
8765	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8764
2916	Q	R	L	L	L	L	E	V	S	W	E	A	L	E	R	A	G	I	D	P		2915
8815	S	L	R	G	S	R	G	S	R	G	G	V	Y	V	A	A	H	G	S	Y	A	8814
2936	S	L	R	G	S	R	G	S	R	G	G	V	Y	V	A	A	H	G	S	Y	A	2935
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P	R	L	L	V	P	E	G	S	E	G	I	L	L	T	G	S	A	D	2955
8884	A	L	A	H	R	A	A	T	G	L	A	A	T	S	A	T	S	A	G			8883
2956	S	D	P																			

Figure 23 cont.





15522 TCC TAC CCG TCG CAC CTC TGC GGC AAC ATC CAG TTC GTG TTC AAC ACC GGC 15581  
 459 C Y R S A H L F Q G K I E F V L S N S G 478  
 15582 CAC ATC CAG ACC ATC CTC AAC CCG CCA GGC AAC CCG AAC GGC TTC ATG ACC GGT GGC 15641  
 479 H I Q S I L N P P G N P K A R F H T C A 498  
 15642 GAT CCG CCG GGT CAC CCG CTC GGC TCG CAG GAA AAC GGC ACC AAC GAT CCG CAC TTC TCG 15701  
 499 D R P G D P V A W Q E N A T K H A D S W 518  
 15702 TCG CTC CAC TCG CAA ACC TCG CTC GGC CAG GGT GGC GGC CAG CTC GAA AAC GGC CCG ACC 15761  
 519 W L H W Q S W L G E R A G E L E K A P T 538  
 15762 CCG CTC CAG AAC CCG TAC TAC GGT GGC CAG GCA TTC CCG GGC ACC TAC GGT CAC CAG 15821  
 539 R L G N R A Y A A G E A S P G T Y V H E 558  
 15822 CCG TCA CTTGACGCGCGCGTGGCCACTCCGCGACGCGCAGCGTGTGATTC 15872  
 559 R 560

Figure 23 cont.

## INTERNATIONAL SEARCH REPORT

International application No. PCT/US96/20119	
<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC(6) : C12P 7/62; C12N 9/88, 15/63; C07H 21/04 US CL : 435/135, 232, 320.1; 536/23.2 According to International Patent Classification (IPC) or to both national classification and IPC	
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) U.S. : 435/135, 232, 320.1; 536/23.2 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched NONE	
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Please See Extra Sheet.	
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>	
Category*	Citation of document, with indication, where appropriate, of the relevant passages
X --- Y	US 5,371,002 A (DENNIS ET AL.) 06 December 1994, column 9, line 65 through column 10, line 62 and Figure 9.
Y	FRASER, M.J. The Baculovirus-Infected Insect Cell as a Eukaryotic Gene Expression System. Curr. Top. Microbiol. Immunol. August 1992, Vol. 158, pages 131-172, especially pages 145-158.
Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.	
Special categories of cited documents: * "A" documents defining the general state of the art which is not considered to be of particular relevance * "E" earlier documents published on or after the international filing date * "L" documents which may throw doubt on priority claim(s) or which are cited to establish the publication date of another citation or other special reason (as specified) * "O" documents referring to an oral disclosure, use, exhibition or other means * "P" documents published prior to the international filing date but later than the priority date claimed * "T" later documents published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention * "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone * "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art * "Z" document member of the same patent family	
Date of mailing of the international search report <b>15 APR 1997</b>	
Name and mailing address of the ISA/US International Bureau of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer KAWAI LAU Telephone No. (703) 308-0196
Date of actual completion of the international search 03 APRIL 1997	

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US96/20119

## C (Continuation), DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JOSHI, A.N. et al. Construction, Expression, and Characterization of a Mutated Animal Fatty Acid Synthase Deficient in the Dehydratase Function. J. Biol. Chem. 25 October 1993, Vol. 268, No. 30, pages 22508-22513, see whole document.	13-21, 28, 29-32, 40-49, 51
X	US 5,229,279 A (PEOPLES ET AL.) 20 July 1993, column 15, lines 43-68.	37, 58
Y	DONADIO, S. et al. Organization of the Enzymatic Domains in the Multifunctional Polyketide Synthase Involved in Erythromycin Formation in <i>Saccharopolyspora erythraea</i> . Gene. 01 February 1992, Vol. 111, No. 1, pages 51-60, see whole document.	30, 31, 40-49
Y	BEVITT, D.J. et al. 6-Deoxyerythronolide-B Synthase 2 from <i>Saccharopolyspora erythraea</i> . Eur. J. Biochem. February 1992, Vol. 204, No. 1, pages 39-49, see whole document.	30, 31, 40-49
X	Database Medline on Dialog, US National Library of Medicine, (Bethesda, MD, USA), No. 95187308, HAN, L. et al. 'Cloning and Characterization of Polyketide Synthase Genes for Jadomycin B Biosynthesis in <i>Streptomyces venezuelae</i> ISP5230', abstract, Microbiol., Vol. 140, No. 12, pages 3379-89, December 1994, see entire abstract.	53

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US96/20119

## B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

AFS; STN-indices Bioscience Patents  
search terms: PHA, PHB, polyhydroxyalkanoate, polyhydroxybutyrate, polyketide, fatty acid synthase, dehydratase,  
*Alcaligenes eutrophus*, *Saccharopolyspora erythraea*, *S. venezuelae*